Program-Level Variation in Cadet Outcomes at the National Guard Youth ChalleNGe Program

Lauren Malone and Jen Atkin

December 2016

CNA
ANALYSIS & SOLUTIONS

This document contains the best opinion of CNA at the time of issue.

It does not necessarily represent the opinion of the sponsor.

Distribution

DISTRIBUTION STATEMENT A. Approved for public release: distribution unlimited.
PUBLIC RELEASE. 12/23/2016

Other requests for this document shall be referred to CNA Document Center at inquiries@cna.org.

Photography Credit: Cadre speak to cadets of Class 2016-1. Photograph provided by Steven Friederich

Approved by: **December 2016**

Jeffery Peterson – Research Team Leader
Fleet and Operational Manpower Team
Resource Analysis Division

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
12-2016	Final	

4. TITLE AND SUBTITLE
Program-Level Variation in Cadet Outcomes at the National Guard Youth ChalleNGe Program

5a. CONTRACT NUMBER
N00014-16-D-5003

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
N/A

6. AUTHOR(S)
Lauren Malone, Jen Atkin

5d. PROJECT NUMBER
N/A

5e. TASK NUMBER
A30520

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Center for Naval Analyses
3003 Washington Blvd
Arlington, VA 22201

8. PERFORMING ORGANIZATION REPORT NUMBER
DRM-2016-U-014458-Final

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Deputy Assistant Secretary of Defense
For Reserve Affairs (R)
1500 Defense Pentagon
Washington, DC 20301-1500

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

DISTRIBUTION STATEMENT A. Approved for public release: distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This study analyzes the relationship between general philosophies/practices at each of seven National Guard Youth ChalleNGe Program sites and cadet outcomes. It also considers the relationship between these outcomes and population demographics. The research question originated from an earlier CNA study in which we found that cadets' final cognitive skills, final noncognitive skills, and probability of completing the ChalleNGe program were affected by which of the seven sites they attended. We hypothesized that there were two potential explanations for these significant site effects: that the populations served by each program differed significantly in demographic and socioeconomic terms and/or that the sites differed significantly in their overarching philosophies and program administration. Our findings suggest that programmatic differences in philosophy, general practices, and classroom instruction are responsible for the role of the specific program attended in determining cadets' final outcomes.

15. SUBJECT TERMS
ChalleNGe, Cadet, Noncognitive, Cognitive, Youth, At-Risk, Test of Adult Basic Education, TABE, GED, Completion, High School

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			Knowledge Center/Robert Richards
U	U	U	SAR	70	**19b. TELEPHONE NUMBER** *(include area code)* 703-824-2123

Abstract

This study analyzes the relationship between general philosophies/practices at each of seven National Guard Youth ChalleNGe Program sites and cadet outcomes. It also considers the relationship between these outcomes and population demographics. The research question originated from an earlier CNA study in which we found that cadets' final cognitive skills, final noncognitive skills, and probability of completing the ChalleNGe program were affected by which of the seven sites they attended. We hypothesized that there were two potential explanations for these significant site effects: that the populations served by each program differed significantly in demographic and socioeconomic terms and/or that the sites differed significantly in their overarching philosophies and program administration. Our findings suggest that programmatic differences in philosophy, general practices, and classroom instruction are responsible for the role of the specific program attended in determining cadets' final outcomes.

This page intentionally left blank.

Executive Summary

In a report released in June 2016, CNA found that National Guard Youth ChalleNGe Program (ChalleNGe) cadets were more likely to experience cognitive and noncognitive gains during their 22 weeks at the program if they attended certain ChalleNGe programs [1]. That is, even after controlling for cadets' gender, age, and initial abilities (both cognitive and noncognitive), cadets' final cognitive and noncognitive skills—as well as their likelihood of completing the program—varied significantly across the seven sites included in the study:

1. Grizzly Youth Academy in California (CA)
2. Fort Gordon Youth ChalleNGe Academy in Georgia (GA)
3. Lincoln's ChalleNGe Academy in Illinois (IL)
4. Youth ChalleNGe Program–Gillis Long in Louisiana (LA)
5. Freestate ChalleNGe Academy in Maryland (MD)
6. Washington Youth Academy in Washington (WA)
7. Wisconsin ChalleNGe Academy in Wisconsin (WI)

We hypothesized that those "site effects" were likely a reflection of variation in (a) the socioeconomic and demographic characteristics of the populations served by each program, (b) the sites' philosophies and practices, or (c) some combination of the two. In this follow-on effort, we attempt to identify whether the programmatic differences or the population differences, or both, were responsible for the importance of the specific site attended in predicting cadets' final outcomes.

In this vein, we revisited each of the seven sites to collect information on how the programs differ, from both programmatic and pedagogical perspectives. Thus, during our visits, we interviewed each director about his program's philosophies and practices. Among other questions, we asked each director what his main goals are for the cadets, how he would characterize the average cadet on arrival at ChalleNGe (in terms of academics, socioeconomics, and self-discipline), whether he thinks the development of cognitive or noncognitive skills is more important, and how often the cadets communicate with family while at the program. In addition, we observed teachers and cadets in the classroom. We noted, for example, which pedagogical

techniques were used in the classroom, how the teachers tried to motivate their students, the different ways in which the cadets participated in class, and whether attempts were made to improve cadets' noncognitive skills.

In addition to these site visits, we used two other data sources in our analysis. First, we used American Community Survey (ACS) data to characterize the populations served by each of the seven programs. This provided socioeconomic and demographic information not for cadets' families, but for those households located in the primary counties that each program's cadets hail from. Second, we used the dataset established by our previous study, which contains the FY15 cadets' basic demographic information, their scores on the Test of Adult Basic Education (TABE), indicators of their noncognitive skills (collected via survey), and whether they completed the program. This dataset, combined with the ACS data and our site-visit information, allows us to determine which population and programmatic characteristics are statistically significantly correlated with cadet outcomes. Due to the nature of our observations and the fact that the ACS data are for the surrounding population, *not* the program's cadets, our resultant variables for programmatic and population characteristics are *average* values. That is, they are the same for all cadets at a given program. As a result, the variables were too correlated with each other for more than a few variables to be included in any given regression. We are therefore unable to analyze the *independent* effects of the population and programmatic differences, but we can determine which of these variables are correlated with cadets' final cognitive scores, their final noncognitive skills, and their likelihood of completing ChalleNGe.

We found, overall, that cadets' noncognitive skills tended to be higher by program completion when:

- Cadets visit the program for an onsite orientation before intake day
- Instructors are given more than minimal guidance regarding how to deliver the curriculum and run their classroom
- Directors cited anything *other than* behavior and discipline as their program's biggest challenge
- We observed a higher rate of classroom participation
- Teachers corrected behavior by reminding students of the rules
- Overall classroom behavior was good
- At least some one-on-one instruction was provided
- A variety of pedagogical techniques were used in a class period

Cadets' cognitive skills were higher by program completion and they were more likely to complete ChalleNGe when:

- Programs measure success as change in cadet attitude
- There are more hours of academic instruction daily
- Students are grouped into classes based on their ages or TABE scores (as opposed to alphabetically or randomly)
- Families are more supportive

Our analysis of the correlation between cadets' socioeconomic/demographic characteristics and their outcomes was largely inconclusive, leading us to conclude that differences in the populations served by the various ChalleNGe programs are likely *not* the primary explanation for significant site effects found in our previous report. Conversely, it appears that the variation in site philosophies and practices (both in and out of the classroom) is the primary driver of the site differences we previously found—namely, that even after controlling for the cadets' basic demographic characteristics and incoming skills, the particular site they attended was a significant predictor of their final cognitive skills, final noncognitive skills, and likelihood of program completion.

Overall, our findings suggest that the following actions could be helpful in improving cadet outcomes: visiting the program for an onsite orientation before intake day, providing instructor guidance (as opposed to leaving lesson plan design and classroom management completely up to the instructors), increasing the focus on discipline and cadet attitude, increasing cadet participation in the classroom, increasing the number of pedagogical methods used in the classroom, organizing classes by age or TABE, and making efforts to increase family buy-in and involvement. Although we do not suggest that any broad-sweeping changes be made to the ChalleNGe program based on our findings, we do recommend that any programs struggling with cadet outcomes consider making some of these adjustments.

This page intentionally left blank.

Contents

This page intentionally left blank.

List of Figures

This page intentionally left blank.

List of Tables

This page intentionally left blank.

Glossary

ACS	American Community Survey
CA	California
ChalleNGe	National Guard Youth ChalleNGe Program
GA	Georgia
GED	Tests of General Educational Development
HiSET	High School Equivalency Test
IL	Illinois
LA	Louisiana
MD	Maryland
PRAP	Post-Residential Action Plan
TABE	Test of Adult Basic Education
TAC	Transition, Acceptance, and Commitment
TRADOC	U.S. Army Training and Doctrine Command
WA	Washington
WI	Wisconsin

This page intentionally left blank.

Introduction

In a recently released CNA study, we analyzed the cognitive and noncognitive growth of cadets participating in the National Guard Youth ChalleNGe Program (ChalleNGe) [1]. Specifically, we surveyed cadets at seven ChalleNGe sites, at both the beginning and the end of the program, and analyzed the determinants of the following:

- Growth in cadets' noncognitive skills (including grit, locus of control, and confidence in their math and science abilities)
- Growth in cadets' cognitive skills, as measured by performance on the Test of Adult Basic Education (TABE)
- Whether cadets completed the ChalleNGe program

The results were surprising. We found significant site effects in each of our estimations; that is, the specific site a cadet attended was an important determinant of whether he or she completed the program as well as of the cognitive and noncognitive growth he or she experienced. Our interpretation of these significant site effects was that they reflected either (a) differences in the sites' philosophies and practices, (b) differences in the populations the sites serve (from a demographic and socioeconomic standpoint), or (c) some combination of these two effects. In this follow-on effort, we aim to determine which of these effects is at play.

To accomplish this goal, we revisited each of the seven sites where we initially conducted the cadet surveys: the Grizzly Youth Academy (California), Fort Gordon Youth ChalleNGe Academy (Georgia), Lincoln's ChalleNGe Academy (Illinois), Youth ChalleNGe Program–Gillis Long (Louisiana), the Freestate ChalleNGe Academy (Maryland), the Washington Youth Academy (Washington), and the Wisconsin ChalleNGe Academy (Wisconsin).[1] At each site, we interviewed the program directors regarding their programs' core philosophies and practices. In addition, we observed a number of classes at each site to determine how the curriculum was delivered and how behavior and other issues were handled. These observations, together with the

[1] In the remainder of this research memorandum, we refer to these seven ChalleNGe program sites using standard two-letter state abbreviations for California (CA), Georgia (GA), Illinois (IL), Louisiana (LA), Maryland (MD), Washington (WA), and Wisconsin (WI).

directors' inputs, were used to identify the primary differences in how the seven sites execute the ChalleNGe program. We then turned to the American Community Survey (ACS) to identify socioeconomic and demographic differences in the areas from which the programs primarily recruit. Our goal is to identify which of these factors (site differences and/or socioeconomic and demographic differences) are correlated with differences in the ChalleNGe cadets' performance across the seven sites.

The remainder of this report is organized as follows. The next section contains a brief overview of the ChalleNGe program and its objectives. Then we provide detailed information on our data and methodology, including descriptions of the information gathered in our director interviews and classroom observations as well as of the ACS data we use. The next three sections summarize the primary differences we found—first in terms of site philosophies and practices, then in terms of what we observed in the classrooms, and finally in terms of the local populations' characteristics. In the final section, we present our conclusions.

National Guard Youth ChalleNGe Program

The National Guard Youth Challenge Program (ChalleNGe) is designed to provide a second chance to high school dropouts and support for those at risk of dropping out (students who have earned far fewer credits than expected). Eligible youth are ages 16 to 18. ChalleNGe consists of a 22-week residential program and a 12-month post-residential mentoring component. During the 12-month follow-up, cadets and their mentors are asked to report back to the program about whether the cadet is employed, in school, or serving in the military. The goal of ChalleNGe is to help "young people improve their self-esteem, self-confidence, life skills, education levels, and employment potential" [2]. Currently, there are 37 ChalleNGe locations in 27 states and the territory of Puerto Rico. These sites are funded jointly by the Department of Defense and the states. The National Guard Bureau is responsible for management and oversight of ChalleNGe. That said, each site is given broad discretion in how it structures its program.

As a result, the academic goals of the ChalleNGe sites vary. Some seek to have cadets pass the Tests of General Educational Development (GED) or the High School Equivalency Test (HiSET), while others award alternative high school diplomas. Some ChalleNGe sites provide credit recovery so that cadets can earn high school credits and return to their original high schools after completing the program. There also are some ChalleNGe sites that are equivalent to high schools and award state-certified high school diplomas. In many cases, sites offer more than one of these options.

ChalleNGe has a quasi-military structure: participants live in barracks, wear military-style uniforms, and perform activities typically associated with military training (e.g., marching, drills, and physical training). Participation, however, is voluntary. Although participants are referred to as cadets, they have no subsequent requirement for military service.

The academic program is administered in a manner similar to that found in a traditional high school setting. Teachers are given curricular guidelines regarding the topics they must cover but otherwise have some latitude regarding classroom management and pedagogical methods. In addition to providing an academic program, ChalleNGe seeks to instill life skills in the cadets. Toward that end, the core

values of ChalleNGe are honor, courage, and commitment. The program has eight core components:

1. Leadership/followership
2. Responsible citizenship
3. Service to community
4. Life-coping skills
5. Physical fitness
6. Health and hygiene
7. Job skills
8. Academic excellence

All of these core values and components focus cadets toward the changes needed to become productive citizens on completion of ChalleNGe.

Data and Methodology

As discussed earlier, our primary objective was to determine *why* there were significant site effects in our previous analysis, indicating that the specific site a cadet attended was important in predicting his or her ChalleNGe success (whether measured by final noncognitive abilities, final cognitive abilities, or the likelihood of completing the program). These effects indicate that, when holding cadet characteristics at their average values, there is significant variation in ChalleNGe success depending on which site a cadet attended. In that earlier work [1], we hypothesized that this was most likely the result of the combined effects of two factors:

1. The sites differ substantially from each other, not only because they offer different educational options (high school diploma, GED, credit recovery), but also because the site directors are given significant leeway in deciding how to run their programs.

2. The sites serve very different cadet populations, in terms of both demographics and socioeconomics, as we illustrated in our previous report using ACS data. These differences could influence not only the noncognitive and cognitive skill levels with which cadets arrive at ChalleNGe, but also how receptive they are to the program's efforts to influence their skills and outlooks.

In this follow-on effort, we have collected data on both of these possible factors. We revisited each of the seven sites included in the previous analysis to conduct interviews with each of the directors and observe classrooms. This allowed us to gather information on how the directors' philosophies and the sites' corresponding implementation of the ChalleNGe program differ. In addition, the seven sites provided information on the primary recruiting areas for their program—that is, the primary counties their cadets hail from. We then consulted the ACS and extracted data for these counties to determine any demographic and/or socioeconomic differences across the programs' recruiting areas. The ACS and site-visit data were added to our existing dataset that contains the cadets' noncognitive survey responses and cognitive abilities (TABE scores) as well as other cadet characteristics (e.g., age and gender). This allowed us to identify which demographic/socioeconomic characteristics and site-level differences are correlated with program outcomes, even

after controlling for the cadets' individual characteristics. (For a complete discussion of our preexisting dataset and its contents, see our 2016 report [1].)

In the remainder of this section, we discuss our data limitations, the specific questions we asked the program directors, what we hoped to gain from conducting the classroom observations, and the variables we extracted from the ACS. We then summarize how we used these data sources to determine whether site-level differences in philosophies and practices (and/or population differences) are correlated with ChalleNGe outcomes.

Data limitations

A few important data limitations prevented us from conducting the ideal analysis to answer the study questions. First, the cognitive, noncognitive, and ChalleNGe completion measures (the outcome variables) are for cadets in the FY15 population—when we conducted our onsite surveys. The cadets observed in the past few months, of course, are *different* cadets. Our analysis therefore hinges on the assumption that there was not significant change over time in the directors' philosophies, the overall program administration at each site, or the way the sites (and their teachers) operate their classrooms. That is, we assume that, on average, the prevalence of certain site-wide philosophies/practices and pedagogical techniques, motivational methods, or behavioral corrections within the classroom is constant over time.

Second, in order to confidently separate the impacts of demographic and socioeconomic characteristics, overall site philosophies, and classroom practices, we would need cadet-level data on all of these characteristics. We would need to know, for example, the income and education levels of each cadet's parents or guardians, the number of times each cadet was visited by family members or received a letter from home, the number of times each cadet exhibited bad classroom behavior (and how it was dealt with), and the pedagogical techniques that each cadet was exposed to in each classroom. Collecting such detailed information was not feasible for a number of reasons, including time and budgetary constraints, the requisite consent processes to collect such information (which, for many of these cadets, who are minors, involve parents/guardians), and the fact that such detailed and intrusive data-collection processes would likely interfere with the primary program objectives and daily operations.

In the absence of cadet-level data on demographic/socioeconomic characteristics or their personal experiences at ChalleNGe, we relied on ACS data to provide the *average* characteristics of people from the same counties that the majority of the programs' cadets hail from. Similarly, we collected information from program directors on their overall philosophies and practices, asking *them*, for example, how often communicating with parents is a challenge because English is their second

language. The directors' responses, therefore, can be viewed as averages for their programs, and those *same* average values are assigned to *all* cadets. Finally, from our classroom observations, we aren't capturing *which* cadets received one-on-one instruction or punitive measures in a certain class period; we only capture the *average* number of classrooms in which these methods were used. All cadets at a given program are thus assigned the same *average* values for all classroom observation variables, not individual values. This greatly reduces the variation in our sample and weakens the precision of our estimates. In addition, we are unable to answer questions of causality with average-level data—whether certain program practices or pedagogical techniques *caused* cadets to have better (or worse) outcomes—but we can determine, on average, what characteristics are *correlated* with positive outcomes for cadets.

Finally, we were only at each site for a few days, thus observing a snapshot of classes. While we note what we observed, we are not in a position to say whether anything we observed was a systemic problem (or strength). That is, we do not know if the behaviors and practices we observed were, in fact, representative of an average week at each program, or whether what we observed differed drastically from what we would have observed had we visited during any *other* week. Our analysis relies on the assumption that what we observed is similar to what we would have observed at any other time, but we can neither confirm nor deny this.

Data collected via site visits

There were two main components of our site-visit data collection efforts: inputs provided by the program directors and data collected via classroom observations. At each site, we spent one to two hours with the director, discussing such topics as program focus, the director's assessment of the average characteristics and abilities of their populations, their relative focus on cognitive versus noncognitive skills, program administration, and the degree of cadet contact with family while at ChalleNGe. The full list of interview questions can be found in Appendix A. Here we provide a subset of the questions we asked:

- What is the mission of the ChalleNGe program, and how do you see the mission of your program?
- What are the goals your program hopes to achieve with each cadet?
- What is the biggest challenge your program faces as it relates to the cadets? Is it low academic skills, controlling their behavior, getting sufficient support from mentors and family members, or something else?
- When cadets first arrive at ChalleNGe, what are they most in need of?

- What have you found to be the most effective way to improve noncognitive skills? Cognitive skills?
- What is the role of military staff at your program?
- Do the cadets come to your program for orientation prior to intake day?
- How often do cadets have the opportunity to speak with their parents/ guardian on the phone while at ChalleNGe?

Our synthesis of the directors' responses to these (and other) questions allowed us to determine the extent to which the programs differ in their implementation of ChalleNGe's core components. Similar insights were gained from our classroom observations. Based on the limited amount of time we were at each program, we worked with program staff to develop a schedule that would enable us to observe as many classrooms as possible. In every case, we put a primary emphasis on observing core curricular classes (e.g., math, language arts), and less emphasis on elective classes since these varied by program. In Table 1, we display the number of classes (and their corresponding subjects) we observed at each site, as well as the program week (out of 22) during which we conducted our visit.

Table 1. Number of classes observed (total and by subject) and program week of our visit, by ChalleNGe site

Site	No. of classes observed	Pro-gram week	Number of classes by subject					
			English	Math	Science	Social studies	Life skills	Other
CA	12	10	3	3	0	3	3	0
GA	6	5	2	1	2	1	0	0
IL	10	12	3	3	2	2	0	0
LA	8	21	0[a]	0[a]	0[a]	2[a]	0[a]	6[a]
MD	4	12	1	2	0	1	0	0
WA	14	9	2	2	4	2	4	0
WI	16	5	2	3	4	5	2	0

Source: CNA classroom observations.

[a] Because it was week 21, our observations in LA were substantially different from all the other programs. The cadets were no longer broken up by TABE levels for subject-specific learning; they were grouped into "homerooms" and were largely working individually.

In the classrooms, we noted the subject being taught, the gender mix, the number of students, the pedagogical techniques used by the teacher, the behavior modification techniques used, and whether the students were largely engaged/paying attention or distracted. In addition, we noted whether the teacher had control of the classroom, what efforts were made to motivate students, whether computers were used in the classroom (and whether computers were used by the teacher and/or the students),

student participation, and whether any efforts were made to improve the cadets' noncognitive skills (overtly or subtly). Appendix B contains the data collection form that was used in our classroom observations. Our primary objective in collecting this information was to determine how the seven sites differ in how they deliver the curriculum and manage their classrooms. Any characteristics with noticeable differences across the sites were considered in our correlation analysis—where we determined which program variables were associated with ChalleNGe outcomes.

Data collected from the ACS

To determine whether substantial differences exist in cadets' demographic or socioeconomic characteristics, we relied on the directors' knowledge of where their programs primarily recruit from and asked them to identify the counties or metro areas that are home to the majority of their cadets.[2] We then used the ACS's summary tables to identify the average characteristics of people living in these areas. Cadets, of course, will not be a random sampling of each area's overall population; it is therefore possible that cadets' characteristics (and their families' characteristics) will not align with those of the overall population. Not all ChalleNGe programs, however, collect sociodemographic or family characteristics on their individual cadets. Consequently, we are left to assume that areas with overall different populations will also produce cadets with different characteristics.

From the ACS summary tables, we extracted information on the 16-year-old and older population, as well as on the 15- to 19-year-old population. For those 16 and older, we used ACS data on population size, gender, racial/ethnic diversity, employment status, employment industry, and income range. For the 15- to 19-year-old population, the ACS data are more limited; they are largely focused on school enrollment, fertility, and labor force participation. The complete tables containing the average values for all ACS variables we considered (and how they differ across the seven programs) can be found in Appendix C.

We used these data to summarize the socioeconomic and demographic characteristics of the populations in the counties or regions that the majority of each program's cadets come from. We also ultimately created *one* "socioeconomic factor" from all these characteristics; this served as a measure of overall demographic and socioeconomic "disadvantage." The higher the overall socioeconomic factor, the more

[2] This information was collected in support of our previous research effort, summarized in [1]. We are reusing the same data here.

"disadvantaged" the cadets in that state. Although the highest possible value of the factor is 8, the highest we observed was 6. The socioeconomic factor values follow:

- CA = 2
- GA = 5
- IL = 3
- LA = 6
- MD = 2
- WA = 1
- WI = 4

Appendix D contains specific information on how we created this factor.

Correlation analysis

In the case of both the ACS variables and the site-level characteristics, we ultimately have one average value for each site. We know, for example, the average employment rate for the 16 and older population in the recruiting areas, the average number of teachers who used a certain pedagogical method (e.g., lecturing, small group work), or the average number of classrooms in which the students were engaged (as opposed to being distracted). From an analytical perspective, there are two significant drawbacks from having our data organized in this way. The first applies primarily to the ACS data since, although it is unrealistic to have anything *other* than average values for the site-level variables, the socioeconomic information *could* be collected for each cadet. For these variables, we know only the average socioeconomic information from each cadet's home area, but we do not know how each cadet compares with those average values. For example, using the average income level for a cadet's home county may either overstate or understate the actual income level in the cadet's household—that cadet's parents could feasibly make significantly more or significantly less than the county average.

Second, for both the ACS variables as well as the site-level characteristics, every cadet at a particular program will have the same value for those variables because they are *program*-level, not *cadet*-level, variables. This increases the standard errors, thus decreasing the significance levels of these variables, creating instances in which the site-level average is insignificantly correlated with outcomes but the cadet-level values might have been significantly correlated with outcomes. Thus, we may fail to detect some significant relationships between these variables and our outcomes of interest (noncognitive skills, cognitive skills, program completion). In addition, we are limited in the number of these program-level characteristics that can be included in any given estimation—since many of the classroom observation and director input variables are highly correlated. Whether a teacher has control of the classroom, for example, will likely be correlated both with his or her use of behavior modification techniques and with whether the students are engaged and paying attention. It was for this reason that we decided to limit our estimations to correlation analysis and include only a limited number of variables in each estimation equation. As a result, we are unable to simultaneously control for classroom observation variables,

directors' inputs, and ACS characteristics in one estimation equation. This means that we cannot determine whether there are *particular* characteristics that are most predictive of cadet outcomes, after taking all other site and population characteristics into account. That is, we simply are not analytically able to take "all other characteristics into account." We *can*, however, simultaneously control for the average level of socioeconomic disadvantage (using our socioeconomic factor) and either classroom observations or directors' inputs. This, at a minimum, allows us to determine whether site differences are primarily being driven by socioeconomic and demographic differences or by differences in the program directors' driving philosophies or how classes are conducted.

In the case of the classroom observations, we were not able to observe every cadet in every classroom. Because of time and funding constraints, we were limited in the number of days we could spend at each program, so during that time we both interviewed the director and observed as many classrooms as possible. It is therefore conceivable that the average experience we observed is different from the average experience for all cadets at that program (if, for example, we observed a non-representative subset of classrooms). Our findings *will* illustrate, however, program-wide characteristics (whether related to socioeconomics/demographics or to program philosophies/administration) that tend to be correlated with our outcomes of interest, *after* controlling for cadets' individual characteristics. Additional analysis with more individual-level data would be necessary to more precisely identify the socioeconomic/demographic or site-level predictors of noncognitive skills, TABE scores, or ChalleNGe completion.

Site Differences in Philosophy and Functionality

We had two main objectives for each of our site visits, both in the vein of collecting sufficient information to determine whether the seven sites in our study vary in terms of their underlying philosophies and their execution of the ChalleNGe program. First, we interviewed each program director. We then observed a number of classes. We review the findings from our director interviews in this section and our findings from the classroom observations in the next.

In our interviews, we first asked the directors a number of questions related to their cadet population, including questions about the cadets' characteristics on arrival, their biggest challenges as related to the cadets, and how frequently they have problems communicating with cadets (or their families) because English is not a primary language. Table 2 summarizes the directors' responses to these questions. As the table shows, the only questions on which the directors were largely in consensus were those regarding the cadets' self-discipline and immediate needs when they first arrive at the program. All but one of the directors indicated that self-discipline is low when the cadets first arrive—noting that, in most cases, that's the driving factor that brought them to ChalleNGe. Relatedly, all but the LA director said that their greatest immediate need is structure and discipline (the LA director said it was supplies and family support).

Table 2. Characterization of each program's cadet population, according to the seven ChalleNGe program directors

Questions posed to program directors	ChalleNGe site
How would you define your average cadet's academic abilities on arrival?	
Low academic ability	GA, IL, LA, MD
Average for their age or better	CA, WA, WI
How would you define your average cadet's socioeconomic status on arrival?	
Low socioeconomic status	CA, GA, IL, LA, WA
Average socioeconomic or better	MD, WI
How would you define your average cadet's self-discipline on arrival?	
Low/problematic	CA, GA, IL, LA, MD, WA

Questions posed to program directors	ChalleNGe site
How would you define your average cadet's family support on arrival?	
Low	IL, LA, WI
Medium	CA, GA, MD
High	WA
What is the biggest challenge as it relates to the cadets? (multiple choice)	
Low academic ability	None
Behavior/lack of respect	GA, IL, LA
Family/mentor support	CA, WI
Other	MD, WA
How often does your program have difficulty communicating with cadets or families because English is not their primary language?	
A lot	CA, IL, WA
Some	MD
Rare	GA, LA, WI
What do cadets need most when they first arrive?	
Structure/discipline	CA, GA, IL, MD, WA, WI
Family support/supplies	LA

Source: Tabulations of information collected from interviews with program directors.

The responses to all other questions, all of which relate to the characteristics of the cadets they serve and/or their families, often varied across the seven ChalleNGe programs. This suggests that there are differences in cadets at each of the seven sites (whether because of differences in academic abilities, socioeconomic status, family support, or the prevalence of non-English-speaking parents)—differences that we can reasonably expect to influence the likelihood of ChalleNGe success.

We found it particularly interesting that there was such variation in directors' responses to what their biggest challenge is as it relates to the cadets. The directors were asked to choose from the following: (1) the cadets arrive at ChalleNGe with low academic skills, (2) the cadets have challenges controlling their behavior and/or have a lack of respect for authority, (3) the program struggles with getting mentors, as well as the cadets' own families, to support the cadets, and (4) something else. While *none* of the directors replied that low academic skills are the biggest challenge, the GA, IL, and LA directors said it is behavior/lack of respect, the CA and WI directors said it is lack of family and mentor support, and the WA and MD directors chose "something else." The WA director indicated that his biggest challenge is emotional resiliency, and the MD director said it is finding (and keeping) competent staff. Although the IL director stated that his biggest challenge is behavior/lack of respect, discussions revealed that this behavioral challenge is actually *also* related to a staffing problem. Specifically, the program struggles to find (and retain) the

necessary cadre (the staff members that oversee the day-to-day lives of the cadets). This is largely because the position offers no health benefits (due to a 2014 change in IL law). As a result, the program is unable to find what it would define as "good" cadre staff (i.e., those who will be fully committed to their roles and impose the necessary behavioral corrections on cadets), and the program suffers from severe behavioral and disciplinary problems. The level of variation in the directors' responses to this question reveals that the seven ChalleNGe programs have *distinct* challenges. Therefore, it is not entirely surprising that there was a site-level impact on the probability of ChalleNGe success revealed by the estimations in our previous report—that is, that the probability of being successful at ChalleNGe depended in part on *which* of the seven sites the cadet attended.

We also learned that there are significant differences in the seven sites' underlying philosophies and program administration, which we summarize in Table 3. First, there were differences in how the directors described what they aim to achieve with their cadets. When asked, for example, whether they thought the development of cognitive or noncognitive skills in cadets is more important, only the CA, GA, and WI directors said "noncognitive," whereas all other directors indicated that they are equally important. Similarly, there were differences in the directors' descriptions of their primary goals for each cadet. The CA, GA, LA, WA, and WI directors all spoke about noncognitive growth, while the MD and IL directors spoke about career goal development. We also asked the directors how they define success when determining if a cadet has had a successful experience at ChalleNGe. Some directors mentioned long-term changes (e.g., the ability to take responsibility for their actions or an improvement in overall maturity), some mentioned academic growth, and others focused on an attitude improvement (the MD director noted all three).

Table 3. Characterizations of underlying site philosophies and program administration, according to the seven ChalleNGe directors

Questions posed to program directors	Program site
Does your program consider cognitive or noncognitive skills to be more important (or neither)?	
Cognitive	None
Noncognitive	CA, GA, WI
Neither (i.e., both are equally important)	IL, LA, MD, WA
What are your primary goals for each cadet?	
Noncognitive growth	CA, GA, LA, WA, WI
Career goal development	MD, IL
How does your program define success?	
Long-term changes	GA, IL, MD, WI
Academic growth	CA, MD, WA
Attitude improvement	LA, MD

Questions posed to program directors	Program site
Do cadets come on site for orientation prior to intake day?	
Yes	CA, GA, MD, WA, WI
No	IL, LA
Does your program use TABE as an admission criterion?	
Yes	WA
No	CA, GA, IL, LA, MD, WI
Do cadets have the same subjects at the same time every day?	
Yes	CA, GA, MD, WI
No	IL, WA
How many hours of academic instruction do cadets have each day?	
Less than 5 hours	IL, WI
More than 5 hours	CA, GA, LA, MD, WA
How involved are the director and the lead instructor in providing guidance to instructors?	
Little involvement	GA, IL, MD
Significant involvement	CA, LA, WA, WI
Are cadets required to write letters?	
To their mentors	GA, IL, LA, WA, WI
To both parents and mentors	CA
No letters required	MD
Are families permitted to take cadets off campus during any visitations?	
Yes	IL, MD
No	CA, GA, LA, WA, WI
What education options are available?	
GED	CA, IL, LA, MD, GA, WA, WI
HSDG	CA, GA
Credit recovery	MD, WA

Source: Tabulations of information collected from interviews with program directors.

There also were notable differences in the directors' explanations of how they administer their programs, such as their orientation processes, how they make admission decisions, their academic program, and the degree to which cadets interact with family members while at ChalleNGe. The programs' orientation processes differ, most notably in how involved the orientation process is and whether the cadets visit the ChalleNGe site before intake day (as part of orientation). Cadets *do* visit the CA, GA, MD, WA, and WI sites prior to intake, but they do not visit the IL or LA sites. If this increased exposure to the ChalleNGe program before intake

day increases the commitment level of those cadets who still choose to attend, then site visits might affect the cadets' overall success (and, in particular, graduation rate). The scope of the orientation process also varied from site to site; the CA and WI directors described the most involved processes. In CA, all cadets are interviewed and required to tour the program. In addition, CA has a Transition, Acceptance, and Commitment (TAC) program, which is highly recommended for all cadets, but especially those questioning their commitment. During TAC, they return to the site for another four to six hours and experience "A Day in the Life of a Cadet," in an effort to provide a better understanding of what the program is about. In WI, a key part of the orientation process is an interactive event during which applicants meet some cadets in the current class (who take them on tours), meet the cadre, and talk about what is bringing them to ChalleNGe. While the cadets are going through "A Day in the Life of a Cadet," there is a separate question-and-answer session for parents that gives them the opportunity to ask questions of some cadets. These were notably different from the processes at some other programs. In LA, for example, the orientation consists of interviews and a 45-minute presentation at National Guard armories throughout the state, two months before the cycle begins. Similarly, in IL, recruiters throughout the state hold orientations in their local armories and speak to the applicants about what to expect at ChalleNGe. Only occasionally will a few cadre members attend, to run applicants through some exercises and give them a taste of ChalleNGe. In fact, the IL director noted that he considers the acclimation phase (the first two weeks, also known as "pre-ChalleNGe") to be a more intense orientation.

The directors also revealed site-level differences in how they conduct the academic portion of their program. Only the WA program, for example, uses the TABE exam as an admission criterion (with a distinct minimum cutoff); the others consider the TABE, in addition to a number of other factors, in making admission decisions, but they do not have a policy that applicants must meet a minimum TABE threshold. The program with a minimum TABE criterion finds it is necessary to get cadets to the high school level, or at least the academic level necessary to pass the GED. In other words, if cadets were admitted below the TABE cutoff, the probability of their success would be low. Conversely, those *without* a minimum TABE policy argue that such policies exclude sections of the population that may be most at need. There is also variation in whether the cadets have the same academic subject at the same time every day. For example, always having math class in the afternoon could serve to disadvantage students who learn better in the morning. If these "morning advantaged" students always have math in the afternoon, their achievable improvements could be lessened.

As Table 3 reveals, there are site differences in the number of hours of academic instruction per day (fewer hours for those sites offering only the GED option) as well as the involvement of the lead instructor in providing guidance to the teachers (less involvement/oversight in GA, IL, and MD). The programs also differ in the amount of communication that occurs between cadets and their mentors and/or parents

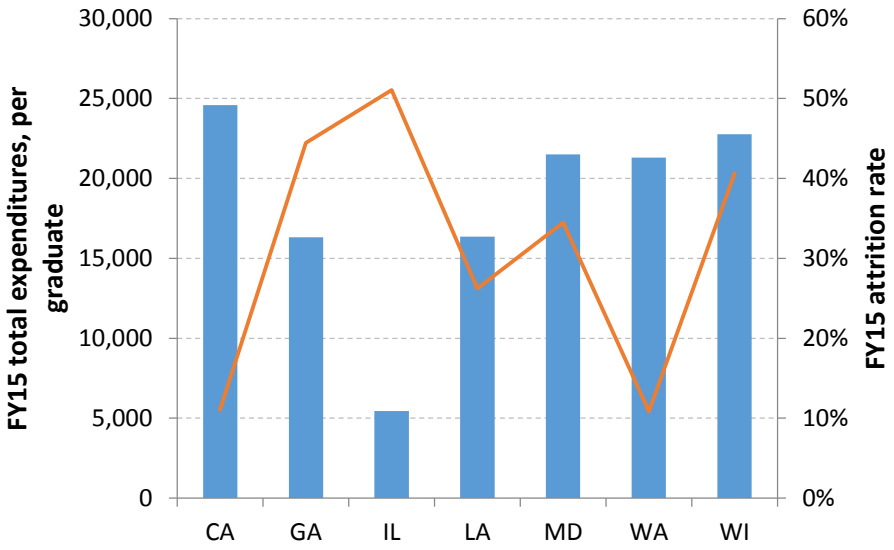
throughout the 22-week residential phase. This communication is typically in the form of letter writing and phone calls. While both can serve to help cadets maintain relationships with their parents and family members as well as to build relationships with their mentors, letter writing can also serve an academic purpose of helping cadets improve their writing skills. Of the seven programs, only CA requires that cadets write letters to their parents. At the other extreme, MD requires no letter writing, but the other five programs all mandate letters to mentors. The cadets are free to write to their family as well, but only the mentor letters are required. Cadets at all programs are allowed weekly phone calls and, in most cases, are free to call whomever they wish in their allotted 5 to 10 minutes. The CA program, however, mandates that the cadets call either their mentor or their parents each week (corresponding to the week that the mentor or parents are *not* receiving a letter from that cadet).

The most striking difference reported in terms of family interactions, however, did not relate to letter writing or phone calls, but rather to family visitation days. At two of the seven programs (IL and MD), the cadets can be taken off campus during these days. At the IL program, there are four Sunday family visitations during the course of the program. The families arrive at 11 a.m. and can take cadets off campus until as late as 6 p.m. Similarly, in MD, there are family days in weeks 7 and 14 during which families may take cadets off campus for up to 10 hours. This is distinct from the other five programs, where family days are restricted to on-campus activities.

Finally, we collected information from the program directors on their sites' annual expenditures. Specifically, they provided their programs' FY15 expenditures (since this was the year of our cadet surveys) in the following broad categories: staff expenditures, stipends/allowances, facilities, furnishings, transportation, dining, supplies, equipment, clothing, recreational equipment, services, communications, medical, security, outreach, and computers and software. The programs' total FY15 expenditures *per graduate* are displayed in Figure 1. In the process of collecting this information, we realized that there are inconsistencies in the way the different ChalleNGe sites compute their expenditures, meaning that the program-specific values are not necessarily comparable. One program noted, for example, that it included in-kind transfers it received in its expenditure values since these transfers were part of the total cost of operating the program in FY15. It is not clear, however, that all programs calculate their expenditures accordingly. In addition, it is worth noting that our presentation of expenditures per graduate may skew some programs' numbers. Specifically, any programs with higher than average attrition will appear to have a higher per cadet cost because the expenditures from the cadets who do not graduate get spread across those who do. Thus, for every cadet who attrites, the cost per graduate will artificially increase; the cost of *any given graduate* hasn't actually risen. For example, a particularly high value for expenditures per graduate could reflect either that the program spends significantly more per graduate than the other

programs *or* that the program experienced significantly higher attrition. The programs' FY15 attrition rates are also displayed in Figure 1.

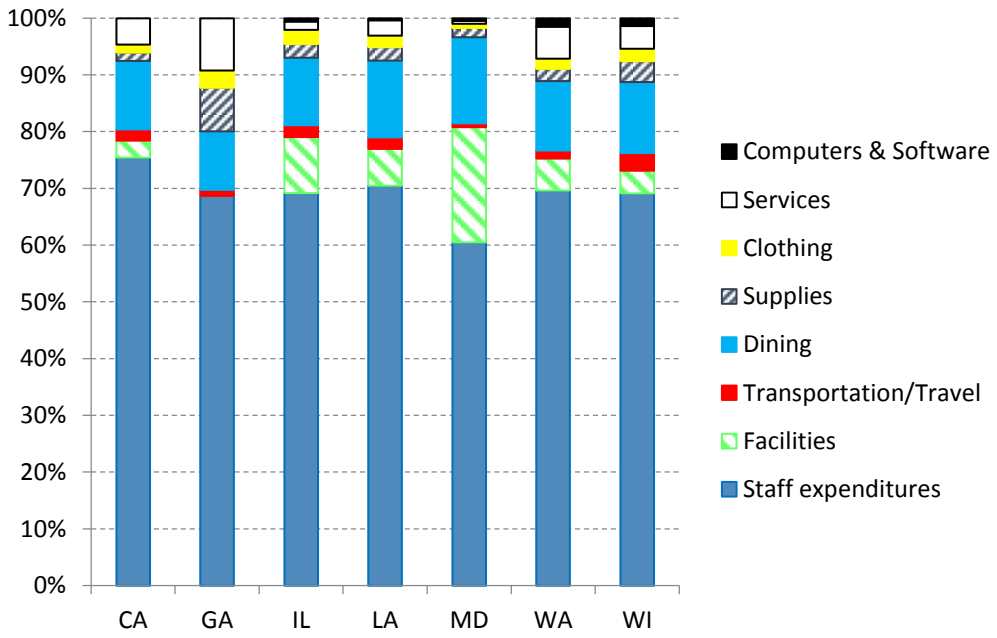
Figure 1. FY15 total expenditures, per graduate



Source: FY15 budget and attrition data provided by the seven ChalleNGe programs.

In Figure 2, we present the percentage distribution of each program's FY15 per-graduate expenditures across a number of different categories. For ease of presentation, those categories that make up only a small percentage of the expenditures are not shown. Our intent is to show how the largest contributors to program expenditures vary across the seven sites. The figure reveals that all seven sites spend 60 percent or more of their total expenditures on staff. Of the remaining categories displayed in Figure 2, those with the most variation across the seven sites are Facilities, Supplies, and Services. In addition, the variation in Computers and Software expenditures is notable, even though it makes up a small percentage of total expenditures (less than 5 percent). What is most striking is that CA and GA have no expenditures in this category, whereas the other seven sites do (this is despite our observation that both locations have computer resources and CA, in particular, had significant resources compared with the other sites). The variations in the sites' expenditures could influence cadets' opportunities and experiences at ChalleNGe, to the extent that the expenditures are reflective of the resources and programs available to them. Because of the possible inconsistencies in how each of the ChalleNGe sites computes its expenditures, however, it is unclear how comparable these numbers really are and thus to what extent they reflect differences in opportunities for the cadets.

Figure 2. FY15 percentage distribution of expenditures by category, for the largest expenditure categories, by ChalleNGe program



Source: Budget data provided by the seven ChalleNGe programs.

[a.] The apparent absence of Facilities costs at the GA program, based on this figure, is because GA's FY15 Facilities Requirements expenditures are listed as $23.25—and that is *before* dividing by the number of FY15 graduates. In response to our inquiry about this, the program director indicated that his facilities expenditures are, in fact, much higher but are incorporated in some other section of the budget information.

A few other site differences that might influence the cadets' experience at ChalleNGe and thus their overall success are not presented in the preceding tables and figures. Some of these differences follow:

- The MD director noted that, in the case of non-English-speaking parents, he has the cadets interpret for the parents. The IL director stated that this sometimes occurs, depending on the subject matter. The other directors, however, specifically noted that they avoid this practice because they believe it is not fair for the cadet to be responsible for translating and because there is no way to validate that the cadet translated accurately (or, potentially, purposefully translated inaccurately).

- Only the CA program requires some of the staff to be members of the military. At that program, all staff members other than teachers and the fiscal staff must have some military affiliation.

- The CA program is also the only program that makes available books to read for pleasure, which likely encourages the development of reading as a habit. There is a program-wide system for checking these books out of the classrooms.

- The classrooms at the MD program are notably sparse. Because the program shares space with U.S. Army Training and Doctrine Command (TRADOC), the space must be kept to TRADOC standards, so the teachers cannot "decorate" or post academic posters or other materials. As a result, the classroom environment is bare, which doesn't provide the most conducive learning environment.

- Relatedly, we observed grimy classroom floors at the GA program. Granted, the cadets track a lot of dirt into the classrooms because of the site's layout and shortage of sidewalks. It was unclear, however, how often the floors are cleaned. This situation might hamper learning (e.g., we observed cadets getting distracted because they wanted to ensure their caps didn't fall on the floor and become soiled).

- The WA program is the only program where cadets can lose a home pass for bad behavior. All other program directors indicated that there was no way to lose a home pass, and some mentioned that this was because their staff members need a break too.

- The IL program is the only program *without* designated mentor-cadet time throughout the course of the program. That is, there are *no* mentor visitation days. The director did note that if a mentor were to show up unannounced, he would never be denied the opportunity to sit and talk with the cadet, but the lack of designated visitations certainly decreases mentor-cadet interactions and the development of that relationship (and the extent to which the mentor can provide encouragement) over the course of the program.

- At the GA program, the credit recovery and high school diploma cadets do not have teacher-directed classrooms. Rather, all of their work is independent and self-directed on computers. We question how effective this construct may be for youth who have left the traditional high school setting and are in need of additional direction.

Having reviewed the site-level differences in program philosophies and administration, we now investigate whether any of these differences are correlated with cadet outcomes. These estimation results are displayed in Table 4 at the end of this section.

In the interest of conserving space, we include the results for multiple different regressions *per outcome* in the table. Specifically, four different regressions were run

per outcome variable, separated by the solid horizontal lines in the table. We were unable to simply include all of these variables in one regression because many of these variables are highly correlated with each other. Thus, if all director-input variables are included in one model, many of them drop out, providing no information as to whether they are correlated with the outcomes or not. We also ran each regression twice—once with a socioeconomic factor variable included and once without. In the table, unbolded results are those we found in the estimation *without* controlling for the socioeconomic factor; bolded results held both when the factor was and was not included. We are therefore most confident in the relationships displayed by the bolded results. We refrain from including a regression with budget information because we are not confident in the comparability of the different sites' budget numbers.

In terms of the cadets' final noncognitive skills (in the first six columns of the table), we find a few relatively consistent relationships (when significant). First, cadets who attended programs that include a visit to the program before intake day had significantly higher noncognitive skills at the end of the program, as did those whose instructors were, on average, given more guidance regarding the content and delivery of their lesson plans. Noncognitive skills were significantly *lower* for those attending programs citing their biggest challenge as behavior (as compared with low academic ability, mentor and family support, or other reasons).

The site philosophy and program administration variables that are correlated with cadets' final cognitive skills and likelihood of completing the ChalleNGe program are often the same. Both of these outcomes are positively correlated with cadet success being measured as a change in attitude (as opposed to long-term growth or academic improvement) and with the number of hours of academic instruction daily. That is, cadets at programs that consider the most important measure of success to be an attitudinal change and that have more hours of academic instruction per day have higher post-TABE scores and are more likely to complete the program. Post-TABE scores were also positively correlated with the programs that cite cadet behavior or lack of respect as their biggest challenge. Cadets attending programs that group cadets into classes based on their age (separating the older cadets from the younger cadets) or their pre-TABE scores (placing cadets of different initial ability levels in different classrooms) were more likely to finish the program and to do so with higher post-TABE scores. We also found that the average level of family support the program receives matters: cadets at programs with overall low levels of family support were less likely to complete ChalleNGe and were more likely to have lower TABE scores. The two unexpected findings were (1) the negative correlation of cadets being required to visit the program prior to intake with post-TABE scores and program completion and (2) the positive correlation between programs' indications that behavior is their biggest challenge with higher post-TABE scores.

Because of the limitations of our data collection, our findings are not sufficiently robust to recommend that certain program philosophies or administrative practices be adopted at all ChalleNGe programs. Our findings *do* suggest, however, that there is a role for these philosophies and administrative practices in determining the likelihood of cadets' success at ChalleNGe (whether measured by their final cognitive skills, their final noncognitive skills, or program completion). The differential site effects found in our previous report, therefore, are likely at least partially the result of the variation in program administration and directors' philosophies across the seven sites.

Table 4. Multivariate analysis of the relationship between site-specific characteristics and ChalleNGe outcomes (noncognitive skills, cognitive skills, and program completion)[a, b, c]

	Final Grit	Final Locus	Final Math Efficacy	Final Science Efficacy	Final Chose $100 Over $50	Final Followed Directions	Post-TABE Overall Battery	Complete Chal-leNGe
Cadets visit before intake	**Insig.**	>0	**>0**	**>0**	>0	**Insig.**	<0	**<0**
Success measure: Change in attitude	**Insig.**	Insig.	**>0**	**<0**	Insig.	**Insig.**	**>0**	**>0**
Program's goals for cadets: Noncognitive growth	**Insig.**	Insig.	>0	<0	Insig.	**Insig.**	**<0**	>0
TABE used in admission decisions	**>0**	Insig.	**Insig.**	**<0**	>0	**Insig.**	Insig.	**Insig.**
Biggest challenge: Behavior	Insig.	**<0**	Insig.	**Insig.**	<0	**Insig.**	**>0**	Insig.
Family support: Low	**<0**	Insig.	**<0**	**>0**	**Insig.**	**Insig.**	**<0**	**<0**
Family support: Medium	**<0**	Insig.	Insig.	**>0**	**Insig.**	**Insig.**	**Insig.**	**<0**
Same subjects at same time each day	**Insig.**	>0	**>0**	**>0**	**Insig.**	**Insig.**	**<0**	**<0**
Hours of academic instruction	**Insig.**	**Insig.**	**>0**	**<0**	**Insig.**	**Insig.**	**>0**	**>0**
Instructors given guidance	**Insig.**	**>0**	**Insig.**	**Insig.**	**>0**	**Insig.**	**<0**	**>0**
Classes organized by age	Insig.	**Insig.**	**>0**	**<0**	<0	**Insig.**	**>0**	**>0**
Classes organized by TABE	**Insig.**	**Insig.**	**>0**	**Insig.**	**<0**	**Insig.**	**>0**	**>0**

Source: Analysis of CNA seven-site cadet survey data and classroom observation data.

[a.] Each row in this table denotes a new regression. Other controls included in each of these regressions are the cadets' initial noncognitive skills, initial TABE scores, gender, and age. Complete regression results are available on request.

[b.] Entries of ">0" or "<0" indicate that the relationship between that variable and the outcome of interest (Final Grit, Final Locus of Control, etc.) is statistically significant at the 10-percent level or better.

[c.] At the suggestion of our peer reviewer, we ran each estimation two ways: with and without a "factor variable" on the right-hand side. This variable, explained in greater detail in Appendix D, allows us to include the program's surrounding areas' socioeconomic characteristics while controlling for the site-specific characteristics in these estimations. Unbolded results held only when the socioeconomic factor was not included. Bolded results indicate that the finding held both before *and* after including this factor. That is, these findings hold whether or not we consider socioeconomic characteristics; they are our most robust findings.

Site Differences: Classroom Observations

In this section, we summarize our classroom observations from our visits to each of the seven ChalleNGe sites. We begin by displaying how the sites differed in terms of their average characteristics and then describe the results of our multivariate analysis where these classroom observation variables were included in our models. More specifically, we first illustrate ways in which the seven sites' average classroom characteristics differ, and then we evaluate whether any of these characteristics are correlated with our metrics of program success. Recall that all of the variables presented here are averages of what we observed (they are "observed averages"); because we observed numerous classrooms at each site, they represent what we observed *on average* at each site. We neither observed *all* classrooms nor observed the same classrooms at numerous points in time, so they do *not* represent the classroom average for the entire site. Table 5, for example displays the average observed prevalence of different pedagogical techniques in the ChalleNGe classrooms, including lecturing, small-group work, one-on-one instruction, and individual work. It also shows the percentage of observed classrooms at each site that used more than one pedagogical technique (of those previously listed) and the average number of techniques used in the observed classrooms. A few notable findings emerge from the table:

- Most observed classroom time was spent lecturing or having the cadets work individually.[3]
- Small-group work and one-on-one instruction were less frequent; we observed no small-group work in GA, IL, or WI and no one-on-one instruction in GA.
- The majority of teachers we observed at all programs used more than one pedagogical method in the course of a classroom observation (in GA it was precisely 50 percent).
- The average number of pedagogical methods used per classroom observation exceeded one at all sites (and exceeded two at all but GA, IL, and LA).

[3] LA is the exception, likely because our visit was close to graduation and the cadets were focused on working on their Post-Residential Action Plans (PRAPs).

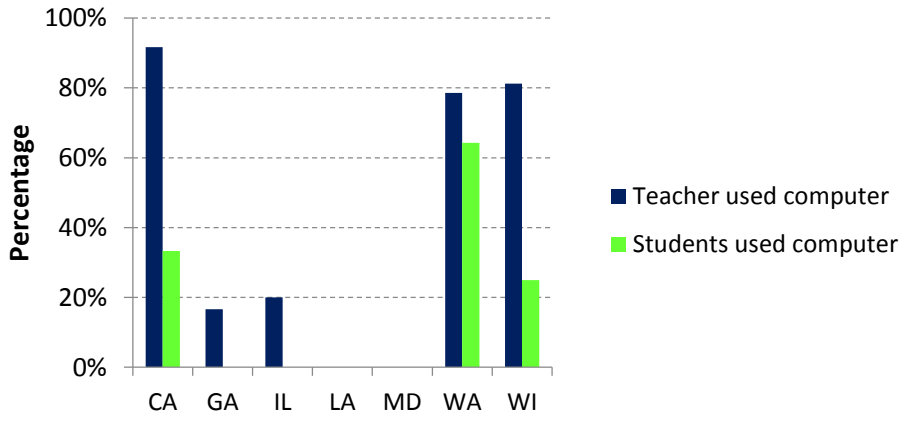
Table 5. Average prevalence of different pedagogical techniques in the classroom, by ChalleNGe site

Site	Lecturing	Small-group work	One-on-one instruction	Individual work	More than 1 pedagogical technique used	Number of pedagogical techniques used
CA	83.3%	41.7%	33.3%	91.7%	83.3%	2.5
IL	100.0%	0.0%	30.0%	70.0%	70.0%	2.0
GA	83.3%	0.0%	0.0%	66.7%	50.0%	1.5
LA	37.5%	12.5%	37.5%	100.0%	62.5%	1.9
MD	100.0%	25.0%	25.0%	100.0%	100.0%	2.5
WA	85.7%	35.7%	64.3%	92.9%	92.9%	2.8
WI	93.8%	0.0%	56.3%	93.8%	87.5%	2.4

Source: CNA classroom observations.

In addition to the teachers' pedagogical techniques, we also took note of whether computers (including Smartboards and other related technologies) were used in the classroom, either by the teacher or by the students. The average prevalence of computer use in the classrooms we observed is displayed in Figure 3. Both teacher and student computer use varied greatly. Teachers used computers in 79 to 92 percent (in increasing order) of WA, WI, and CA observed classrooms, but in only 20 percent of IL observed classrooms, in only 17 percent of GA observed classrooms, and in none of the LA or MD observed classrooms.

Figure 3. Average classroom computer use, by ChalleNGe site



Source: CNA classroom observations.

In Table 6, we present information on students' average participation at each of the sites. The first four columns contain the percentages of classrooms in which we observed students participating via certain methods (e.g., raising their hands; helping each other; behaving well—defined as being quiet, awake, and paying attention; and/or showing respectfulness), while the last column shows the average participation rate (defined as the percentage of students who *actively* participated in class, going above and beyond the minimum required of them). Nearly all students who spoke in class first raised their hands (as opposed to speaking without first being called on by the teachers) at all sites except GA, IL, and LA—where cadets raised their hands to speak in 33.3, 70, and 62.5 percent of the classrooms we observed. The cadets were well behaved in 75 percent or more of classes at all sites except GA and IL, where the cadets were well behaved in only 50 percent of the classrooms we observed.[4] This appears to be correlated with the percentage of classrooms in which the cadets showed respectfulness—lowest in GA at 33.3 percent. Cadets participated by helping each other in a minimum of 0 percent of observed classes in GA and a maximum of 64.3 percent of observed classes in WI. The average participation rate differed across programs as well—with more than 80 percent of observed cadets participating in LA and WI, compared with less than 40 percent in CA, GA, and IL.

Table 6. Average student participation methods, by ChalleNGe site

Site	Majority of students raised hands	Some students helped each other	Majority of students showed good behavior	Majority of students showed respectfulness	Participation rate
CA	100.0%	25.0%	83.3%	100.0%	36.3%
GA	33.3%	0.0%	50.0%	33.3%	35.7%
IL	70.0%	30.0%	50.0%	60.0%	35.5%
LA	62.5%	62.5%	87.5%	87.5%	87.5%
MD	100.0%	25.0%	75.0%	75.0%	45.0%
WA	92.9%	64.3%	92.9%	92.9%	54.1%
WI	100.0%	56.3%	93.8%	100.0%	90.3%

Source: CNA classroom observations.

[4] At the GA, IL, and MD programs, there were a noticeable number of cadets sleeping in the classroom, and not being awakened. That is, they were napping in class and this appeared to be tolerated.

Table 7 contains information on the observed average prevalence of different behavior modification techniques used in the classrooms. Note that these data represent only cases where the instructor attempted to improve bad behavior; we did observe cases where cadets displayed bad behavior that was not addressed in any way by the instructor. The most common technique used to modify students' classroom behavior, at all sites, is to "remind students of the rules." This behavior modification technique was used in 62.5 and 66.7 percent of the classrooms we observed in CA and GA, respectively, and 70 percent or more of the classrooms observed at all other programs. The other commonality across all programs *except* GA is that the second most common technique is to "refocus students." This is not necessarily surprising since these are the techniques that require the least interruption (a teacher can ask for the students' attention or remind them to be silent without deviating significantly from the content being taught). What we found somewhat surprising are the following observations: notable lack of punitive measures in the observed IL classrooms (and very few in the observed WA classrooms); the fact that no teachers whose classrooms we observed in CA, GA, MD, or WI used peers to positively influence other students; and the fact that, at *all* programs, positive reinforcements were used as a behavior modification technique at most in 37.5 percent of the observed classrooms (in IL). To the extent that some of these behavior modification techniques are more impactful than others, such variation could suggest that poor behavior is more likely to be a frequent classroom distraction at some programs than others.

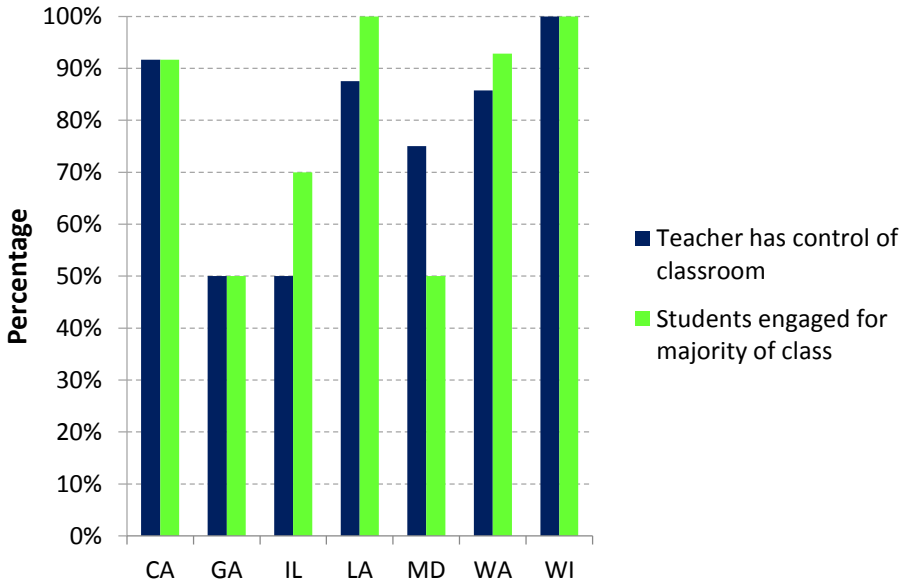
Table 7. Average prevalence of different behavior modification techniques, by ChalleNGe site

Site	Refocus students	Punitive measures	Positive reinforcements	Use peers to positively influence students	Remind students of the rules
CA	50.0%	33.3%	16.7%	0.0%	83.3%
GA	0.0%	50.0%	0.0%	0.0%	66.7%
IL	70.0%	0.0%	20.0%	10.0%	90.0%
LA	62.5%	25.0%	37.5%	12.5%	62.5%
MD	75.0%	50.0%	25.0%	0.0%	100.0%
WA	57.1%	7.1%	35.7%	14.3%	71.4%
WI	81.3%	25.0%	18.8%	0.0%	100.0%

Source: CNA classroom observations.

In addition to the types of behavior modification techniques used, we also took note of, overall, whether the teacher had control of the classroom and whether the students were engaged and paying attention for most of the class period.[5] The average observed differences, by program, are shown in Figure 4. The most striking patterns in the data are for the combination of the teacher having control *and* students being engaged in 100 percent of the classrooms we observed in WI and in over 80 percent of the classrooms in CA, LA, and WA. Conversely, the teacher had control in only 50 percent of the observed classrooms in GA and IL and 75 percent of the observed classrooms in MD. These three programs also had notably lower student engagement in the observed classrooms: 70 percent in IL and 50 percent in GA and MD. We would expect that such differences would correlate with student outcomes—a supposition that will be tested later in this section.

Figure 4. Teacher control of classroom and student engagement, by ChalleNGe site



Source: CNA classroom observations.

[5] Because these are difficult to quantify or measure, the researchers recording the observations discussed in advance what would qualify as a teacher having control of the classroom and the students being engaged/paying attention. Recognize, however, that there is still some subjectivity in how these characteristics are defined and recorded.

The last two classroom "characteristics" we observed were the methods used to increase (or maintain) student motivation and the focus on noncognitive skills in the classroom. The average observed prevalence of different motivation techniques and noncognitive skills are shown in Table 8 and Table 9, respectively. More than one motivation technique was used, on average, in all observed classrooms in CA, IL, MD, and WI, but only 92.9 percent of the time in WA and 50 percent of the time in GA and IL. Using various techniques to keep the cadets motivated is clearly a common practice since we observed it in at least half of the classrooms at all sites. The average number of motivation methods used per observed classroom ranges from a low of 1.7 in GA to a high of 3.4 in WI. The other striking commonality across programs (all with the exception of GA) is the design of lessons to encourage participation; in GA, 0 percent of the lessons we observed were designed to encourage participation, a striking polarity. Differences emerge across all programs, however, in the degree to which teachers vary instructional techniques, call on students who have not raised their hands, and make real-life connections of the subject matter. Of particular note is that instructional techniques were varied in only 12.5 percent of classrooms observed in LA (although this is likely due to the fact that we visited close to the end of the program, when students were working on their PRAPs). In addition, observed teachers in WI, IL, and MD were much more likely to call on students who hadn't raised their hands *and* were also more likely to make real-life connections to the material. One hundred percent of the teachers we observed in GA made real-life connections to the material.

Table 8. Average prevalence of different techniques to increase student motivation, by ChalleNGe site

Site	Instruction techniques were varied	Teacher called on students without raised hands	Lesson designed to encourage participation	Teacher made real-life connections	More than one motivation method used	Number of motivation methods used
CA	100.0%	33.3%	91.7%	41.7%	100.0%	2.7
IL	60.0%	60.0%	100.0%	90.0%	100.0%	3.1
GA	50.0%	16.7%	0.0%	100.0%	50.0%	1.7
LA	12.5%	37.5%	100.0%	50.0%	50.0%	2.0
MD	75.0%	50.0%	100.0%	75.0%	100.0%	3.0
WA	92.9%	35.7%	85.7%	42.9%	92.9%	2.6
WI	93.8%	68.8%	100.0%	81.3%	100.0%	3.4

Source: CNA classroom observations.

Our findings regarding the average focus on noncognitive skills in the classrooms are shown in Table 9. Starting at the rightmost column of the table, there is noticeable variation in the average number of noncognitive skills that are addressed in an

observed class period, ranging from 0.5 in GA to 2.5 in WI and MD. Similarly, we observed at least one noncognitive skill addressed in only 33.3 percent of the GA classrooms and 50 percent of the IL classrooms, compared with 75 percent of the LA and MD classrooms, and over 80 percent of the CA, WI, and WA classrooms (with WA teachers addressing at least one noncognitive skill in 100 percent of our observed classrooms). In terms of the specific noncognitive skills addressed, there is much variation across the sites. The sites at which we observed 50 percent or more classrooms addressing each noncognitive skill are the following:

- CA and WA for organizational skills
- WA, IL, LA, and MD for study skills
- WI for time management
- CA for self-advocacy
- MD for conflict resolution
- WI for perseverance
- WI, LA, and MD for discipline/respect

Also notable is the fact that self-advocacy and conflict resolution were, at a number of sites, addressed in *none* of the classes we observed.

Table 9. Average focus on noncognitive skills during the class period, by ChalleNGe program

	Site						
Noncognitive skill	**CA**	**IL**	**GA**	**LA**	**MD**	**WA**	**WI**
Organizational skills (%)	50.0	10.0	16.7	12.5	25.0	57.1	12.5
Study skills (%)	50.0	50.0	33.3	62.5	75.0	50.0	31.3
Time management (%)	0.0	10.0	0.0	0.0	25.0	28.6	50.0
Self-advocacy (%)	50.0	0.0	0.0	0.0	0,0	0.0	43.8
Conflict resolution (%)	16.7	0.0	0.0	0.0	50.0	0.0	12.5
Perseverance (%)	33.3	0.0	0.0	12.5	25.0	14.3	50.0
Discipline/respect (%)	33.3	20.0	0.0	50.0	50.0	28.6	50.0
Any noncognitive skill (%)	83.3	50.0	33.3	75.0	75.0	100.0	87.5
Number of noncognitive skills	2.2	0.9	0.5	1.4	2.5	1.8	2.5

Source: CNA classroom observations.

Our estimation results regarding the relationship between the classroom observation variables and a number of ChalleNGe outcomes are displayed in Table 10. In the interest of conserving space, we include the results for multiple different regressions *per outcome* in that table, as we did in the previous section. Six different regressions were run per outcome variable, separated by the solid horizontal lines in the table. We were again unable to include all classroom observation variables in one regression because many of these variables are highly correlated with each other.

Thus, if all classroom observation variables are included in one model, many of them drop out, providing no information as to whether they are correlated with the outcomes. In addition, we ran two versions of each regression—one where we included the socioeconomic factor variable and one where we did not. Unbolded results in the table are those that held only when the socioeconomic factor was *not* included; bolded results are more robust and held regardless of whether the socioeconomic factor was included.

In terms of the cadets' final noncognitive skills (in the first six columns of the table), we find that they are often positively correlated with the classroom participation rate, efforts made to remind students of the rules, and good classroom behavior. The pedagogical technique most frequently correlated with the cadets' final noncognitive skills is one-on-one instruction, but most final noncognitive skills also tend to be higher when a variety of pedagogical methods are used within a class period. We found it surprising that the relationship between classroom computer use and noncognitive skills varies: teacher computer use is positively correlated with cadets' final science efficacy and their ability to delay gratification (choosing $100 in six months rather than $50 today); student computer use is negatively correlated with science efficacy (but positively correlated with other skills).

The relationships between our classroom observation variables and the cadets' final cognitive skills, as measured by their overall post-TABE score, also vary. Some methods of behavioral correction are positively correlated with TABE scores (refocusing students and punitive measures), whereas reminding students of the rules is negatively correlated. Similarly, the prevalence of lecturing and one-on-one instruction in the classroom is negatively correlated with TABE scores, whereas small-group work is positively correlated with the final scores. We find similar variation in the correlation of the classroom observation variables with a cadet's likelihood of completing ChalleNGe.

As discussed in the data limitation section, our findings are not sufficiently robust to recommend that certain teaching practices be emphasized (or deemphasized) in all classrooms. What our findings *do* suggest, however, is that there is a role for classroom practices in determining the likelihood of cadets' success at ChalleNGe (whether measured by their final cognitive skills, their final noncognitive skills, or program completion). As a result, the differential site effects found in our previous report are likely at least partially due to the variation in classroom practices across the seven sites.

Table 10. Multivariate analysis of the relationship between classroom observations and ChalleNGe outcomes (noncognitive skills, cognitive skills, and program completion)[a, b, c]

	Final Grit	Final Locus	Final Math Efficacy	Final Science Efficacy	Final Chose $100 Over $50	Final Followed Directions	Post-TABE Overall Battery	Complete ChalleNGe
Participation rate	**>0**	Insig.	Insig.	>0	**>0**	**Insig.**	**<0**	<0
Behavioral correction: Refocus students	Insig.	Insig.	Insig.	**<0**	**<0**	**Insig.**	**>0**	**>0**
Behavioral correction: Punitive measures	**Insig.**	**Insig.**	**>0**	**Insig.**	**<0**	**Insig.**	**>0**	**>0**
Behavioral correction: Remind students of the rules	**Insig.**	**Insig.**	**Insig.**	**>0**	**>0**	**Insig.**	**<0**	**<0**
Number of motivational methods	**Insig.**	**>0**	**<0**	Insig.	>0	**Insig.**	**<0**	**<0**
Good classroom behavior	**>0**	**>0**	**>0**	**<0**	**Insig.**	**Insig.**	**<0**	**>0**
Lecturing	**Insig.**	Insig.	Insig.	**>0**	Insig.	**Insig.**	**<0**	**<0**
Small-group work	**Insig.**	Insig.	**>0**	<0	Insig.	**Insig.**	>0	>0
One-on-one instruction	**>0**	**>0**	**Insig.**	<0	**>0**	**Insig.**	**<0**	**Insig.**
Individual work	**Insig.**	**Insig.**	**Insig.**	>0	Insig.	**Insig.**	Insig.	**Insig.**
Number of pedagogical methods	**>0**	**>0**	>0	**<0**	>0	**Insig.**	**<0**	**>0**
Teacher computer use	**Insig.**	**Insig.**	**Insig.**	**>0**	**>0**	**Insig.**	**<0**	**<0**
Student computer use	**>0**	Insig.	**>0**	**<0**	**Insig.**	**Insig.**	**>0**	**>0**
At least one noncognitive skill addressed	**>0**	**>0**	>0	**<0**	>0	**Insig.**	**<0**	**>0**

Source: Analysis of CNA seven-site cadet survey data and classroom observation data.

[a.] Each horizontal line in this table denotes a new regression. Other controls in each of these regressions include the cadets' initial noncognitive skills, initial TABE scores, gender, and age. Complete regression results are available on request.

[b.] Entries of ">0" or "<0" indicate that the relationship between that variable and the outcome of interest (Final Grit, Final Locus of Control, etc.) is statistically significant at the 10-percent level or better.

[c.] Unbolded results held only when the socioeconomic factor (explained in Appendix D) was not included. Bolded results indicate that the finding held both before and after including this factor. That is, these findings hold whether or not we take socioeconomic characteristics into account; they are our most robust findings.

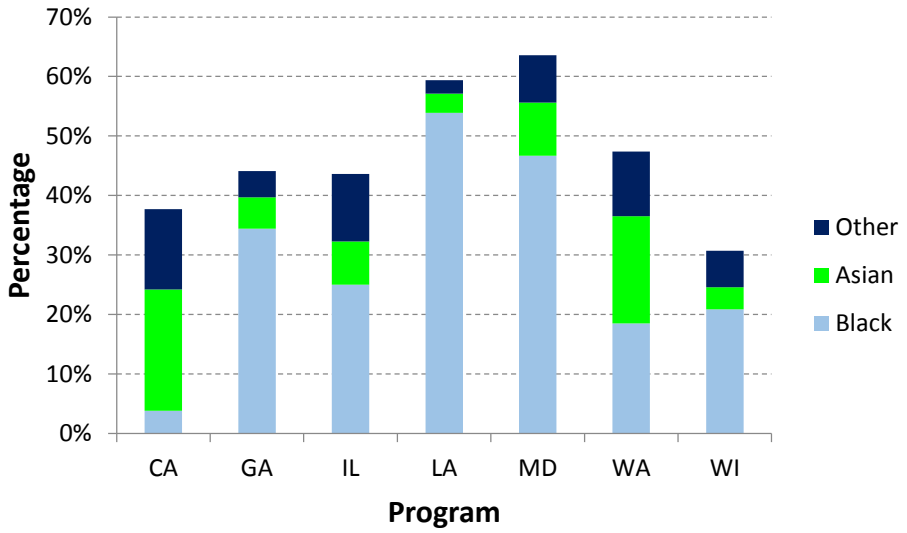
Site Differences: Demographics of the Local Population

We begin this section by presenting the average values of some ACS variables across the seven sites. We focus on those characteristics that were noticeably different from site to site; there were other socioeconomic and demographic characteristics that we pulled from the ACS data, but we do not present those with little difference across the sites. Specifically, the characteristics we illustrate here include racial/ethnic makeup of the 16 and older population, industry of employment, and the income distribution.

Figure 5 illustrates the non-white composition of those age 16 and older; specifically, it shows the percentage of the population that identifies as black, Asian, and other. There are a few particularly striking differences. First, there is variation in the percentage of the local population that is a racial minority—ranging from approximately 30 percent for the WI program to nearly 64 percent for the MD program. In addition, the minority composition varies across the seven sites. In GA, IL, LA, MD, and WI, blacks constitute more than half of the minority population. Conversely, in CA and WA, most minorities identify as Asian or other.

In Figure 6, we display the percentage of the local population that identifies as Hispanic. Once again, there are notable differences across the programs: roughly 35 percent of the local population is Hispanic at the CA and GA programs, and roughly 25 percent is Hispanic in IL and WA. The MD program's surrounding population is only 15 percent Hispanic, and the LA program's surrounding population is only 4.5 percent Hispanic. Such differences suggest that there will likely be differences in the cadets' cultures and norms across the seven ChalleNGe sites, which could influence the likelihood of achieving success at ChalleNGe.

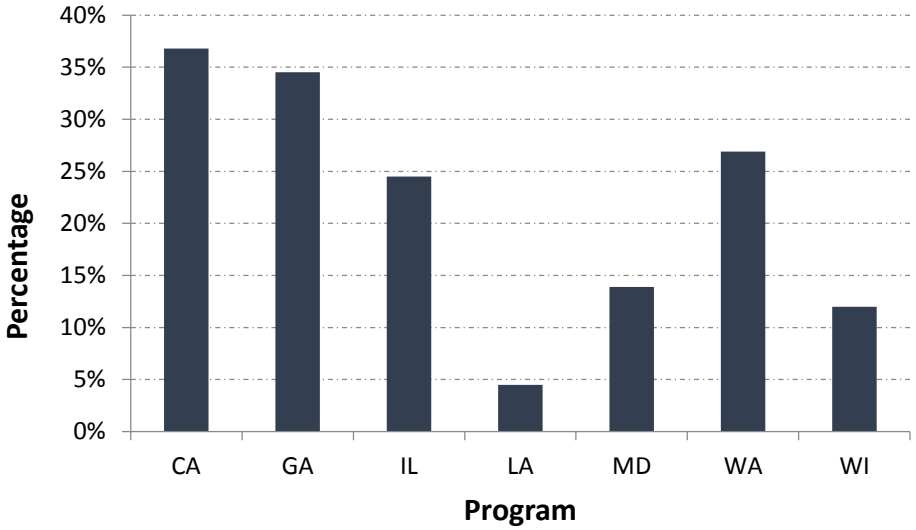
Figure 5. Minority composition of the local population age 16 and older[a]



Source: U.S. Census Bureau, 2010-2014 ACS 5-Year Estimates, Table DP03: Selected Economic Characteristics.

[a.] The "local population" for each ChalleNGe site is defined as the primary counties and metropolitan areas that the site's cadets hail from.

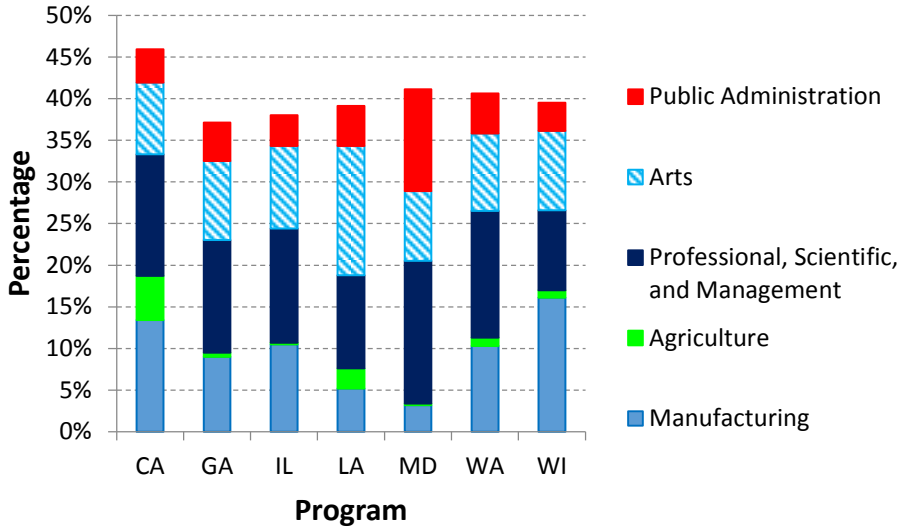
Figure 6. Hispanic percentage of the local population age 16 and older[a]



Note: Source and footnote are the same as for Figure 5.

The other two characteristics across which we noted significant variation were the industry of employment and the income distribution. These are presented in Figure 7 and Figure 8, respectively. The first thing to note is that, at all programs, the combination of the industries shown in Figure 7 accounts for less than 50 percent of the local industries. The remainder falls into the Census's "other" category. There is noticeably more manufacturing in CA and WI, more agriculture in CA, and more public administration in MD. Otherwise, there do not appear to be drastic differences.

Figure 7. Distribution of the local population's industry of employment[a]



Source: U.S. Census Bureau, 2010-2014 ACS 5-Year Estimates, Table DP03: Selected Economic Characteristics.

[a] The "local population" for each ChalleNGe site is defined as the primary counties and metropolitan areas that the site's cadets hail from.

Figure 8 (at the end of this section) shows how the annual income distribution varies across households local to the seven sites. All programs' local populations are most concentrated in the $50,000 to $99,999 range, from 23.4 percent of LA's population to 31 percent of WA's population. Another noticeable difference is that the LA program is in the poorest locale—with nearly 14 percent of households earning less than $10,000 per year and another 21.5 percent earning between $10,000 and $24,999. We might expect the predominant industry of employment and the local income distribution to matter for cadets to the extent that they influence the amount of financial turbulence in their households and thus the academic and extracurricular experiences they had. Later in this section, we will test whether these local demographic and socioeconomic characteristics are correlated with cadets'

outcomes. In these estimations, however, we could not include information both on incomes *and* on the locally prominent industries because the industry and income distributions are highly correlated. We opted for the income distribution—and simplified this to the percentage of households earning $100,000 or more per year.

Before discussing our results regarding which demographic and socioeconomic variables are correlated with our outcomes of interest, it is worth noting the ACS variables that, perhaps surprisingly, varied little from program to program. These include the female share of the population age 16 and older (roughly 50 percent), the percentage of the population in the labor force (ranging only from 63 to 68 percent), the percentage unemployed (roughly 10 percent), the percentage of households with minor children (from 25 to 31 percent), and *all* of the variables specific to the 15- to 19-year-old population.[6] This lack of variation, however, did not prevent us from including them in our models as we tested which variables were significantly correlated with ChalleNGe success. Those demographic and socioeconomic variables not presented in the results that follow were consistently *insignificantly* correlated with our outcome measures, and thus are not included in our results.

Our estimation results regarding the relationship between local demographic or socioeconomic characteristics and a number of ChalleNGe outcomes are displayed in Table 11. We find that cadets at programs with a larger local minority presence, all else equal, have lower noncognitive skills by the end of the program. These cadets, however, have higher cognitive skills, as measured by their overall battery on the post-TABE. The relationship with ChalleNGe completion is mixed—positive for cadets at programs where the local population is a higher Asian percentage but lower for those at programs where the local population is a higher Hispanic percentage. We also find that the local unemployment rate is significantly correlated with ChalleNGe outcomes, all else equal. Cadets at programs where the local unemployment rate is higher have lower noncognitive skills at the end of the program, but their cognitive scores are higher. Finally, the programs' local populations' income distribution is largely insignificant—the one exception is that the percentage of the local population with household incomes greater or equal to $100,000 has a positive and significant relationship with cadets' final science efficacy. Thus, overall, we find only minimal evidence that individual cadets' likelihood of ChalleNGe success (whether measured by noncognitive skills, cognitive skills, or program completion) is related to the average socioeconomic and demographic characteristics in their home areas.

[6] This does not necessarily imply that there is little difference in the characteristics of the 16- to 18-year-olds the ChalleNGe programs are admitting. There are a limited number of ACS variables for the 15- to 19-year-old population—on population size, school enrollment, marital status, fertility, and idleness. Teenagers could be similar on these characteristics but differ in other important ways. Namely, they could come from households with different socioeconomic statuses and could have parents from very different backgrounds.

Table 11. Multivariate analysis of ChalleNGe outcomes (noncognitive skills, cognitive skills, and program completion)[a]

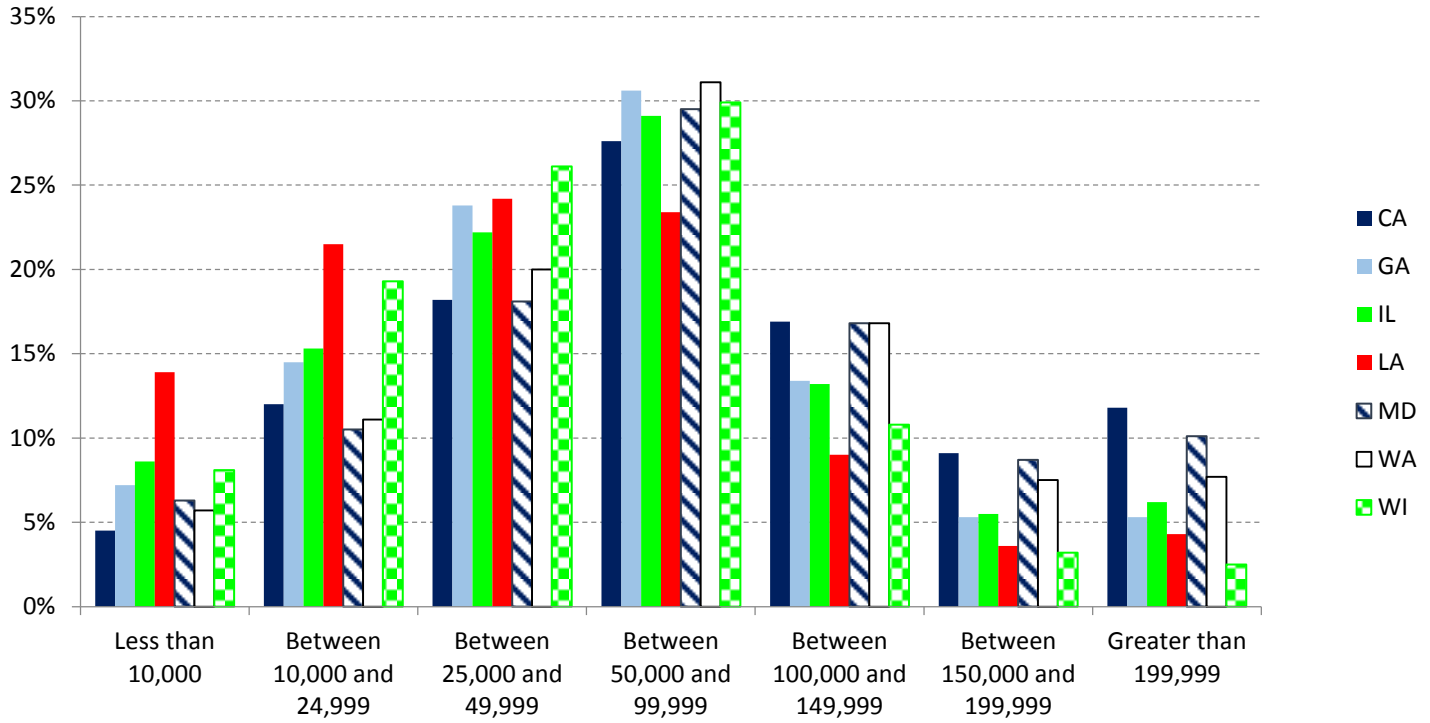
Percentage	Final Grit	Final Locus of Control	Final Math Efficacy	Final Science Efficacy	Final Chose $100 Over $50	Final Followed Directions	Post-TABE Overall Battery	Complete ChalleNGe
Black	Insig.	<0[b]	Insig.	<0	Insig.	Insig.	>0	Insig.
Asian	Insig.	Insig.	Insig.	<0	Insig.	Insig.	>0	>0
Other	Insig.	Insig.	<0	<0	Insig.	Insig.	>0	Insig.
Hispanic	Insig.	Insig.	Insig.	Insig.	Insig.	Insig.	>0	<0
Unemployed	<0	<0	Insig.	Insig.	<0	Insig.	>0	Insig.
Household incomes >=$100,000	Insig.	Insig.	Insig.	>0	Insig.	Insig.	Insig.	Insig.

Source: Analysis of CNA seven-site cadet survey data and ACS data.

[a.] Other controls included in these regressions include the cadets' initial noncognitive skills, initial TABE scores, gender, and age. Complete regression results are available on request.

[b.] Entries of ">0" or "<0" indicate that the relationship between that variable and the outcome of interest (Final Grit, Final Locus of Control, etc.) is statistically significant at the 10-percent level or better.

Figure 8. The local population's annual household income distribution[a]



Source: U.S. Census Bureau, 2010-2014 ACS 5-Year Estimates, Table DP03: Selected Economic Characteristics.

[a.] The "local population" for each ChalleNGe site is defined as the primary counties and metropolitan areas that the site's cadets hail from.

Concluding Remarks

In this report, we provided a detailed analysis of the programmatic and population differences across seven ChalleNGe sites (CA, GA, IL, LA, MD, WA, and WI) in an effort to identify *why* site effects were found to be significant predictors of cadets' final cognitive skills, noncognitive skills, and likelihood of ChalleNGe completion in our previous study. In that study—in which we presented the results of our seven-site cadet survey—we found that, even after taking a cadet's age, gender, and incoming skills (both cognitive and noncognitive) into account, the particular ChalleNGe site a cadet attended mattered for his or her success.

We hypothesized that these site effects were actually capturing differences in the sites' philosophies and practices, differences in the attending populations (demographically and socioeconomically), and/or a combination of these two effects. Since all of the data we collected in this follow-on effort is at the *program* level (and is therefore constant for all cadets attending a specific program), we were not able to econometrically distinguish these effects. That is, we could not simultaneously include the population variables and the programmatic variables in the same estimations because of the high level of correlation between the variables. We were, however, able to create a single socioeconomic factor—which takes higher values for more socioeconomically "disadvantaged" areas—and include this in our classroom observation and site-specific characteristic estimations. As a result, although we *could not* determine which variables are independent predictors of those outcomes, we *could* determine which of these variables were correlated with cadets' outcomes and which of these predictor-cadet outcome relationships are strong enough to sustain the inclusion of our socioeconomic factor in the model.

Our multivariate analysis revealed that a number of programmatic and classroom characteristics are significantly correlated with cadet outcomes. Cadets' final noncognitive skills, for example, are higher at programs where orientation is done on site and before intake day, where more guidance is given to instructors, and where we observed a higher classroom participation rate, teachers more frequently reminding students of the rules, good classroom behavior, more one-on-one instruction, and variety in the pedagogical techniques used within the class period. They are negatively correlated with directors' indication that behavior is their program's biggest challenge. We also found that cadets' final noncognitive skills are higher at programs where we observed greater classroom participation, variety in the number of pedagogical techniques used within a class period (including lecturing,

small-group work, one-on-one instruction, and individual work), incorporation of more time for one-on-one instruction, and teachers reminding misbehaving students of the rules. To effectively incorporate these classroom strategies, program directors will likely need to engage with their counterparts to learn what practices have and have not worked elsewhere. We therefore suggest that the ChalleNGe directors make it a regular practice to share their challenges, lessons learned, and best practices with each other. While the directors can certainly reach out to each other informally, we believe the program as a whole might benefit from the creation of a forum for such communications.

Final cognitive scores and program completion, however, were positively correlated with the programs that measure success (in terms of whether a cadet has had a successful experience) as a change in the cadet's attitude, more hours of academic instruction daily, and classes being grouped based on age or TABE score (as opposed to randomly or alphabetically). Both were negatively correlated with an overall sense of low family support.

Finally, we also found differences across the seven programs in terms of the demographic and socioeconomic characteristics of the populations from which they recruit. The most striking differences were in the percentage of the population that is a racial minority and the minority composition, the primary industries of employment, and households' average incomes. Final noncognitive skills tended to be positively correlated with a higher minority presence, although final cognitive skills were negatively correlated with this measure. Similarly, final noncognitive skills were lower, on average, at programs that recruit from areas with higher unemployment rates, but final cognitive skills were higher at these programs. These inconsistent results suggest to us that the characteristics of the local population are largely not responsible for the differences in cadet outcomes across the seven sites.

In addition, when we included a socioeconomic factor in the estimations of the correlation between classroom and program characteristics with cadet outcomes, most relationships were unchanged. This suggests that site differences are largely due to differences across the seven programs and their classrooms, *not* to differences in the populations they serve. To more confidently arrive at this conclusion, however, we would need individual-level socioeconomic and demographic information for each of the cadets, as opposed to relying on the average characteristics of the population in the programs' primary recruiting areas. That said, we would recommend that any programs struggling with cadet outcomes consider adding an onsite, pre-intake orientation; providing more instructor guidance; increasing the focus on discipline and cadet attitude; increasing cadet participation in the classroom; organizing classes by age or TABE; and making efforts to increase family buy-in and involvement.

Appendix A: Script for Program Director Interviews

Program Focus

What is the mission of the ChalleNGe program?

Does the mission of your program differ from other programs given the population you serve? If so, how is it different and why is it different?

What are the goals your program hopes to achieve with each cadet? Are they the same for all cadets? If not, how do they differ?

In determining if a cadet has had a successful experience at ChalleNGe, how does your program define success?

What is the biggest challenge your program faces as it relates to the cadets? Is it (choose one):

(1) The cadets arrive at ChalleNGe with low academic skills
(2) The cadets have challenges controlling their behavior and/or have a lack of respect for authority
(3) The program struggles with getting mentors, and the cadets' own families, to support the cadets
(4) Other: _____

Your Population

How would you describe your average cadet (when they arrive) in terms of academic abilities? Self-discipline? Socioeconomic status? Family support?

When cadets first arrive at ChalleNGe, what are they most in need of? How does this change, if at all, over the course of the program?

Does your program face any difficulties in communicating with cadets and/or their families due to English not being their primary language?

How would you describe the needs of the population your program serves as compared to the needs of youth in other areas of the country (e.g., our cadets come from households where English is not the primary language spoken)?

Academic vs. Soft Skills

Does your program consider both cognitive and noncognitive improvements to be important for cadets? Is one more important than the other? If so, why?

What have you found to be the most effective way to improve cadets' noncognitive skills?

What about improving their cognitive skills? What is the most effective way to improve cadets' math and reading levels? Does this vary by cadet? If so, what does it depend on?

How Militaristic

How important is the military component of ChalleNGe for these cadets (e.g., drilling/formation, military-like uniforms)? Why?

What is the role of the military staff at your program?

Program Administration

Describe your orientation process. Do cadets come to the program for any length of time or meet with program staff before intake day?

What is the cadets' schedule each day? Does it change over the course of the program?

How much guidance/training is provided to the teaching staff in terms of how to structure their classroom and the most effective teaching approaches for this population?

Connection With Family

How many home passes are granted to each cadet during their time at ChalleNGe? What, if anything, can cause a cadet to lose leave privileges?

How often, if at all, do cadets have the opportunity to speak with their parents/guardian on the phone while at ChalleNGe? At what point in time during the program do those calls take place (i.e., at the end of the 2nd week of the program)?

Are the cadets required to write letters home? If so, how often?

What percentage of your cadets, on average, receive at least one letter from home while they are at ChalleNGe?

Do you have any events that involve family members and/or mentors during the program, with the exception of graduation?

Appendix B: Data Collection Form for Classroom Observations

What subject is being taught?

How many students are in the classroom?

What is the gender mix in the classroom?

Which pedagogical techniques did the teacher use (check all that apply)?

____ Lecturing

____ Small-group work

____ One-on-one instruction

____ Cadets working individually

If they did work in small groups or on their own, did the teacher engage with the groups and/or students as they worked?

Does the teacher have control of classroom?

____ Yes

____ No

What behavior modification techniques did the teacher use (check all that apply)?

____ Refocusing the student(s)

____ Punitive measures; describe: _____

____ Positive reinforcements; describe: _____

____ Use of peers to positively influence fellow students

____ Reminding students of the rules

For the majority of the class, were the students (choose one):

____ Engaged/paying attention?

____ Distracted and doing other things?

What efforts did the teacher make to motivate students (check all that apply)?

____ Instructional techniques were varied; what techniques were used?
_____

____ Teacher called on students who haven't raised their hand

____ Lesson was designed to encourage student participation

____ Teacher explained connection of material to real-life situation(s); how?
_____

Did the lesson include the use of computers? If so, how and to what extent? What computer program or website was used?

To what degree did the students participate in the class?

Approximately how many students participated in the class? _____

In what ways did the students participate (check all that apply)?

____ Students raised their hands

____ Students assisted each other

____ Behavior was generally good

____ Students showed respect toward the teacher/toward each other

Did the lesson include efforts to improve cadets' noncognitive skills? _____ If so, what was done? Which of the following skills were addressed (check all that apply)?

____ Organizational skills

____ Study skills

____ Time management

____ Self-advocacy

____ Conflict resolution strategies

____ Perseverance

____ Discipline/respect

Appendix C: Local Population Characteristics

Table 12. Size and characteristics of the 16 and older population in surrounding areas, by ChalleNGe site[a]

Characteristic	CA	GA	IL	LA	MD	WA	WI
Population age 16 and older	3,080K	4,934K	4,157K	643K	2,001K	2,532K	1,067K
Women age 16 and older	49.9%	52.1%	52.2%	52.5%	52.9%	50.4%	52.0%
Race[b]							
White	66.8%	58.3%	58.6%	42.6%	39.9%	56.7%	72.9%
Black	3.8%	34.4%	25.0%	53.9%	46.7%	18.5%	20.9%
Asian	20.4%	5.3%	7.3%	3.2%	8.9%	18.0%	3.7%
Other	13.5%	4.4%	11.3%	2.3%	8.0%	10.9%	6.1%
Ethnicity							
Hispanic	36.8%	34.5%	24.5%	4.5%	13.9%	26.9%	12.0%
Employment							
In labor force[c]	65.1%	66.7%	66.3%	63.3%	69.9%	67.9%	66.7%
Employed	90.2%	89.3%	88.3%	89.8%	90.7%	91.9%	90.3%
Unemployed	9.8%	10.7%	11.7%	10.2%	9.3%	8.1%	9.7%
Not in labor force	34.9%	33.3%	33.7%	36.7%	30.1%	32.1%	33.3%
Female share of labor force	44.6%	47.9%	48.2%	49.7%	50.3%	45.8%	49.4%
Female share of unemployed	45.9%	49.0%	47.2%	48.7%	48.5%	44.8%	43.7%
Have minor children	30.9%	31.6%	27.8%	24.8%	27.2%	26.8%	29.9%
Civilian employed population age 16 and older	1,803K	2,924K	1,175K	365K	1,263K	1,550K	642K
Industry							
Agriculture, forestry, fishing and hunting, and mining	5.3%	0.5%	0.2%	2.4%	0.2%	1.0%	0.9%
Construction	5.5%	6.4%	4.6%	6.1%	6.4%	5.4%	4.0%

Characteristic	CA	GA	IL	LA	MD	WA	WI
Manufacturing	13.4%	9.0%	10.5%	5.2%	3.2%	10.3%	16.1%
Professional, scientific and mgmt. services	14.6%	13.5%	13.7%	11.2%	17.1%	15.2%	9.6%
Arts, entertainment, recreation and accommodation and food services	8.6%	9.5%	9.9%	15.5%	8.4%	9.3%	9.5%
Public administration	4.0%	4.6%	3.7%	4.8%	12.2%	4.8%	3.4%
Other	48.6%	56.5%	57.4%	54.8%	52.5%	53.9%	56.5%
Income							
Less than $10,000	4.5%	7.2%	8.6%	13.9%	6.3%	5.7%	8.1%
$10,000-$14,999	4.1%	4.6%	4.9%	7.7%	3.7%	3.5%	6.6%
$15,000-$24,999	7.9%	9.9%	10.4%	13.8%	6.8%	7.6%	12.7%
$25,000-$34,999	7.7%	10.0%	9.6%	11.0%	7.2%	8.0%	11.5%
$35,000-$49,999	10.5%	13.8%	12.6%	13.2%	10.9%	12.0%	14.6%
$50,000-$74,999	15.2%	18.3%	17.1%	14.5%	16.9%	17.8%	18.3%
$75,000-$99,999	12.4%	12.3%	12.0%	8.9%	12.6%	13.3%	11.6%
$100,000- $149,999	16.9%	13.4%	13.2%	9.0%	16.8%	16.8%	10.8%
$150,000- $199,999	9.1%	5.3%	5.5%	3.6%	8.7%	7.5%	3.2%
$200,000 or more	11.8%	5.3%	6.2%	4.3%	10.1%	7.7%	2.5%

Source: U.S. Census Bureau, 2010-2014 ACS 5-Year Estimates, Table DP03: Selected Economic Characteristics.

[a.] Percentages may not sum to 100 due to rounding.

[b.] Race distributions are based on the entire population, not just those age 16 and older, due to ACS data availability.

[c.] Throughout this table, *labor force* refers to the civilian labor force. Technically, the labor force also includes the armed forces, but it is not possible to be in the armed forces and unemployed. Thus, we are only interested in statistics that relate to the civilian labor force.

Table 13. Size and characteristics of the 15- to 19-year-old population in surrounding areas, by ChalleNGe site

Characteristic	CA	GA	IL	LA[a]	MD	WA[b]	WI
Population age 15 to 19	281K	450K	343K	50K	166K	186K	94K
Enrolled in school	90.0%	87.4%	88.9%	88.7%	89.1%	87.8%	88.7%
Not enrolled in school	10.0%	12.6%	11.1%	11.3%	10.9%	12.2%	11.3%
Marital status and fertility							
Male	51.5%	51.3%	50.8%	48.5%	51.0%	51.4%	50.8%
Ever married	0.8%	0.9%	0.7%	0.9%	0.7%	0.9%	0.6%
Female	48.5%	48.7%	49.2%	51.5%	49.0%	48.6%	49.2%
Ever married	1.8%	1.3%	1.0%	0.8%	0.9%	1.9%	0.8%
With a birth in the past 12 months	2.1%	0.5%[c]	2.3%	1.8%	2.2%	1.3%	2.4%
Population age 15 to 19 in households	92.3%	93.8%	94.7%	80.7%	89.7%	93.4%	89.8%
In married-couple households	60.5%	54.7%	55.2%	28.7%	47.8%	59.6%	46.1%
Idleness							
Not enrolled in school and not in labor force	4.4%	6.0%	5.8%	5.9%	4.6%	5.2%	4.9%

Source: U.S. Census Bureau, 2010-2014 ACS 5-Year Estimates, Table S0902: Characteristics of Teenagers 15 to 19 Years Old.

a. The Louisiana calculations include statistics for only two of the three counties that were included in Table 1 and that are home to the majority of the site's cadets. This is because ACS data for 15- to 19-year-olds were not available for the Lake Charles metro area.

b. The Washington calculations include statistics for only two of the three counties that were included in Table 1 and that are home to the majority of the site's cadets. This is because ACS data for 15- to 19-year-olds were not available for Franklin County.

c. This statistic is based on data from only two of the three Georgia counties included in the rest of the statistics. This is because ACS does not report a birth rate for the Atlanta metro area.

Appendix D: Creating the "Socioeconomic Factor"

The high level of correlation across various socioeconomic and demographic characteristics (e.g., regions whose predominant industry is manufacturing or agriculture tend to have lower household incomes) prevented us from including our full set of characteristics in one model. In addition, we were unable to include them along with other program-level variables, such as average classroom observations or directors' inputs. Given that our primary research question, however, was whether our previously found site-level effects were being driven by socioeconomics and demographics or by differences in the actual ChalleNGe programs, it was important to be able to simultaneously control for socioeconomic/demographic characteristics *and* program-level differences. We therefore created a "socioeconomic factor," which allowed us to combine an otherwise long list of variables into *one* variable. This variable could then be included in our estimations of cadet outcomes (noncognitive skills, cognitive skills, and program completion). This enables us to determine if the site-specific characteristics and classroom observations are still statistically significantly correlated with cadet outcomes once a measure of the local socioeconomic characteristics is taken into account.

To create this variable, we chose the socioeconomic and demographic characteristics with the most variation across the seven ChalleNGe sites:

- The prevalence of the agriculture, manufacturing, and professional industries
- The racial/ethnic composition
- The unemployment rate
- Income
- Teenagers' family situations

First, each variable was transformed into a binary variable (taking values of 0 or 1). For each input to the socioeconomic factor, we chose cutoffs based on the variation in the ACS data. The aim was to have sufficient variation in each input so that some ChalleNGe sites would have a "1" for that variable and others would have a "0." The cutoff for black and Hispanic, for example, is one-third; any ChalleNGe sites that recruit from areas where the population is at least one-third black take a value of 1 for "black." The selected cutoffs for each variable are displayed in Table 14.

Table 14. The components of the socioeconomic factor and their corresponding cutoff values

Variable	Cutoff value (variable = 1 if value exceeds the cutoff)
Black	1/3
Hispanic	1/3
Unemployment rate	10 percent
Industry share: agriculture	.9 percent
Industry share: manufacturing	9 percent
Industry share: professional	14 percent
Income less than $100,000	75 percent
Teenagers not living in households	10 percent
Teenagers not living in married households	5 percent

We defined these inputs and the overall socioeconomic factor so that a higher value of the factor represents a less advantaged socioeconomic condition. Areas that have a higher minority representation, a higher unemployment rate, a higher prevalence of agriculture or manufacturing, a lower prevalence of professional jobs, more households with income under $100,000, and more teenagers not living in households (whether married or not) will be considered *more* disadvantaged. Thus, in all cases except "industry share: professional," the variable takes a value of 1 if the cutoff is exceeded. For the professional industry, the variable takes a value of -1 if the cutoff is exceeded because an increased prevalence of the professional industry would *decrease* the level of socioeconomic disadvantage. To arrive at our ultimate value of the socioeconomic factor, we add the 0s, 1s, and -1s that are created by the cutoffs. The highest possible level of the factor (the most disadvantaged) is 8. The maximum value we observe in our data is 6. We display the corresponding socioeconomic factor for each ChalleNGe program in Table 15.

Table 15. Socioeconomic factor values, by ChalleNGe program

Program	Socioeconomic factor
CA	2
GA	5
IL	3
LA	6
MD	2
WA	1
WI	4

Source: CNA tabulations.

References

[1] Malone, Lauren D., and Jennifer R. Atkin. 2016. *Cognitive and Noncognitive Improvements Among ChalleNGe Cadets: A Survey of Seven Sites.* CNA. DRM-2016-U-013175-Final.

[2] National Guard Youth Foundation. 2016. "About Youth ChalleNGe." National Guard Youth Foundation. Accessed Sep. 13, 2016. http://www.ngyf.org/about-challenge-2/.

This page intentionally left blank.

CNA

This report was written by CNA's Resource Analysis Division (RAD).

RAD provides analytical services—through empirical research, modeling, and simulation—to help develop, evaluate, and implement policies, practices, and programs that make people, budgets, and assets more effective and efficient. Major areas of research include health research and policy; energy and environment; manpower management; acquisition and cost; infrastructure; and military readiness.

DRM-2016-U-014458-Final-2

CNA is a not-for-profit research organization
that serves the public interest by providing
in-depth analysis and result-oriented solutions
to help government leaders choose
the best course of action
in setting policy and managing operations.

Nobody gets closer—
to the people, to the data, to the problem.

CNA
ANALYSIS & SOLUTIONS

www.cna.org • 703-824-2000

3003 Washington Boulevard, Arlington, VA 22201